



PROCEEDINGS

OF THE
SEVENTH ANNUAL ACQUISITION
RESEARCH SYMPOSIUM
THURSDAY SESSIONS
VOLUME II

**Acquisition Research
Creating Synergy for Informed Change
May 12 - 13, 2010**

Published: 30 April 2010

Approved for public release, distribution unlimited.

Prepared for: Naval Postgraduate School, Monterey, California 93943



ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 30 APR 2010		2. REPORT TYPE		3. DATES COVERED 00-00-2010 to 00-00-2010	
4. TITLE AND SUBTITLE Acquisition Research Creating Synergy for Informed Change. Proceedings of the Seventh Annual Acquisition Research Symposium Thursday Sessions. Volume II				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School, Graduate School of Business & Public Policy, Acquisition Research Program, Monterey, CA, 93943				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 482	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

The research presented at the symposium was supported by the Acquisition Chair of the Graduate School of Business & Public Policy at the Naval Postgraduate School.

To request Defense Acquisition Research or to become a research sponsor, please contact:

NPS Acquisition Research Program
Attn: James B. Greene, RADM, USN, (Ret)
Acquisition Chair
Graduate School of Business and Public Policy
Naval Postgraduate School
555 Dyer Road, Room 332
Monterey, CA 93943-5103
Tel: (831) 656-2092
Fax: (831) 656-2253
E-mail: jbgreene@nps.edu

Copies of the Acquisition Sponsored Research Reports may be printed from our website www.acquisitionresearch.net .



ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL

Preface and Acknowledgements

One needs no powers of prophecy to predict that perilous economic times lie ahead for our nation and the rest of the world. While uncertainty surrounds the global economic crisis—how long it will last, for example, and whether conditions will worsen further before they improve—there can be little doubt that national budgets for defense acquisition will experience considerable pressure and quite possibly decline in the coming years.

Under such conditions, we may be tempted either to accept with fatalism some reduced level of expectations for acquisition outcomes or, alternatively, to embrace trite exhortations to do more with less. Neither of these is, of course, a tenable position for serious scholars of acquisition. Rather, we ought to continue to seek the best possible understanding of acquisition that will lead to the best possible outcomes given available resources. This entails continued work on what Don Kettl has termed the “smart buyer” problem: knowing what to buy, who to buy it from, and how to assess its quality.

Nor can the “how to buy” problem be neglected, as new laws, regulations, and other policies continue to subject acquisition processes to what many see as excessively “bureaucratic” requirements. While these new structures often have worthy goals to increase accountability, transparency, and social equity, their economic costs can’t be ignored. Acquisition scholars should find ways to, paraphrasing Aaron Wildavsky, “speak scientific truth to power” so that policy-makers’ decisions are informed by the best possible research on their costs and benefits.

Currently, DoD investments in acquisition research represent about one one-thousandth of one percent of the total defense budget, yet we believe they have the potential to lead to considerable savings, both in the near and long term. We see strong evidence that this potential is being realized through the products of the Naval Postgraduate School’s Acquisition Research Program.

Our goals remain the same, with recent highlights noted below:

1. Position the ARP in a leadership role to continue to develop the body of knowledge in defense acquisition research
 - Over 450 published works since inception,
 - All completed research is published in full text on the ARP website, www.acquisitionresearch.net, allowing ready access by any and all parties interested in the DoD acquisition process,
 - Sponsoring our 7th Annual Acquisition Research Symposium, the first of which was held in May 2004, draws thought leaders of the DoD acquisition community, academia and industry.
2. Establish acquisition research as an integral part of policy-making for DoD officials. Some processes informed by this research include:
 - Open Architecture implementation practices and policies to include software/hardware reuse repository characteristics such as ontology, search engines, licensing issues and testing requirements; Creation of the concept of Integration Readiness Levels paired with technology readiness levels to create a System Readiness Level scale;
 - Development of logistics resource strategies for selecting among Contracted Logistics Support, Organic Logistics Support or Blended Logistics Support;



- Identification of the scope and causes of bid protests and recommendations for reducing same; and
 - Creation of a database of key contracting workforce demographics for the Army Contracting Command.
3. Create a stream of relevant information concerning the performance of DoD acquisition policies with viable recommendations for continuous process improvement.
 - The body of knowledge on the DoD acquisition process continues to increase by over 140 research products a year.
 - Faculty researchers routinely give multiple presentations, in both national and international fora, featuring their research work—thereby increasing exposure to a broader audience. Typical audiences include the London School of Economics, the Federal Reserve, the Center for Strategic and International Studies, the Western Economics Association International Conference, the International Procurement Conference and the National Contracts Management Association. At a minimum, over 90 presentations of sponsored research results were made in 2009.
 - The *International Journal of Defense Acquisition Management* is the ARP's initiative to promote defense acquisition research in the peer-reviewed literature. All published articles can be freely downloaded from the journal's website. During 2009, three more articles were published in the journal from authors in the US, Australia, and the UK, bringing the total to six since the journal was founded in May 2008. There are also seven articles currently under review. The journal substantially increases the “reach” of our research products.
 4. Prepare the DoD workforce to participate in the continued evolution of the defense acquisition process.
 - The ARP plays a major role in providing a DoD-relevant graduate education program to future DoD officials. Synergy between research conducted and course content delivered enhances both the teaching and learning processes.
 - The number of students engaged in focused acquisition research for their MBA projects continues to grow. These students have the benefit of being able to immediately apply their newly acquired acquisition skills to real-world issues.
 - Student projects on the economics of the shipbuilding industrial base and services contracting are expected to contribute significantly to the body of knowledge and decision-making process for these two very important and timely subjects.
 5. Collaboration among universities, think tanks, industry and government in acquisition research.
 - Over 60 universities/think tanks have participated in the annual Acquisition Research Symposium or the Acquisition Research Program as a result of a focused effort to create a Virtual University Consortium.
 - Emerging collaborative research efforts continue to bring new scholar and practitioner thought to the business issues facing the DoD, as was demonstrated by the large response to our fourth Broad Area Announcement (BAA) in support of the OSD-sponsored acquisition research program. As we write this, our fifth BAA is being prepared for release.



- The International Journal of Defense Acquisition is attracting scholars from the United Kingdom, Canada, Nigeria, Singapore, The Netherlands, the United States and Australia.

We gratefully acknowledge the ongoing support and leadership of our sponsors, whose foresight and vision have assured the continuing success of the Acquisition Research Program:

- Office of the Under Secretary of Defense (Acquisition, Technology & Logistics)
- Program Executive Officer SHIPS
- Commander, Naval Sea Systems Command
- Army Contracting Command, US Army Materiel Command
- Program Manager, Airborne, Maritime and Fixed Station Joint Tactical Radio System
- Program Executive Officer Integrated Warfare Systems
- Office of the Assistant Secretary of the Air Force (Acquisition)
- Office of the Assistant Secretary of the Army (Acquisition, Logistics, & Technology)
- Office of Naval Air Systems Command PMA-290
- Deputy Assistant Secretary of the Navy (Acquisition & Logistics Management)
- Director, Strategic Systems Programs Office
- Director, Defense Business Systems Acquisition Executive, Business Transformation Agency
- Deputy Director, Acquisition Career Management, US Army

We also thank the Naval Postgraduate School Foundation and acknowledge its generous contributions in support of this Symposium.

James B. Greene, Jr.
Rear Admiral, US Navy (Ret.)

Keith F. Snider, PhD
Associate Professor



THIS PAGE INTENTIONALLY LEFT BLANK



The Acquisition Research Program Team

Rear Admiral James B. Greene, Jr. USN (Ret.)—Acquisition Chair, Naval Postgraduate School. RADM Greene develops, implements and oversees the Acquisition Research Program in the Graduate School of Business and Public Policy. He interfaces with DoD, industry and government leaders in acquisition, facilitates graduate student research and conducts guest lectures and seminars. Before serving at NPS, RADM Greene was an independent consultant focusing on Defense Industry business development strategy and execution (for both the public and private sectors), minimizing lifecycle costs through technology applications, alternative financing arrangements for capital-asset procurement, and “red-teaming” corporate proposals for major government procurements.

RADM Greene served as the Assistant Deputy Chief of Naval Operations (Logistics) in the Pentagon from 1991-1995. As Assistant Deputy, he provided oversight, direction and budget development for worldwide US Navy logistics operations. He facilitated depot maintenance, supply chain management, base/station management, environmental programs and logistic advice, and support to the Chief of Naval Operations. Some of his focuses during this time were leading Navy-wide efforts to digitize all technical data (and, therefore, reduce cycle-time) and to develop and implement strategy for procurement of eleven Sealift ships for the rapid deployment forces. He also served as the Senior Military Assistant to the Under Secretary of Defense (Acquisition) from 1987-1990; as such, he advised and counseled the Under Secretary in directing the DoD procurement process.

From 1984-1987, RADM Greene was the Project Manager for the AEGIS project. This was the DoD’s largest acquisition project, with an annual budget in excess of \$5 billion/year. The project provided oversight and management of research, development, design, production, fleet introduction and full lifecycle support of the entire fleet of AEGIS cruisers, destroyers, and weapons systems through more than 2500 industry contracts. From 1980-1984, RADM Greene served as Director, Committee Liaison, Office of Legislative Affairs followed by a tour as the Executive Assistant, to the Assistant Secretary of the Navy (Shipbuilding and Logistics). From 1964-1980, RADM Greene served as a Surface Warfare Officer in various duties, culminating in Command-at-Sea. His assignments included numerous wartime deployments to Vietnam, as well as the Indian Ocean and the Persian Gulf.

RADM Greene received a BS in Electrical Engineering from Brown University in 1964; he earned an MS in Electrical Engineering and an MS in Business Administration from the Naval Postgraduate School in 1973.

RADM Greene received the **2009 Richard W. Hamming Annual Faculty Award for Achievement in Interdisciplinary Activities**. The selection is based on his work in leading and administering the Naval Postgraduate School's Acquisition Research Program.

Dr. Keith F. Snider—Associate Professor of Public Administration and Management in the Graduate School of Business & Public Policy at the Naval Postgraduate School in Monterey, California, where he teaches courses related to defense acquisition management. He also serves as Principal Investigator for the NPS Acquisition Research Program and as Chair of the Acquisition Academic Area.

Snider has a PhD in Public Administration and Public Affairs from Virginia Polytechnic Institute and State University, a Master of Science degree in Operations Research from the Naval Postgraduate School, and a Bachelor of Science degree from the United States Military Academy at West Point. He served as a field artillery officer in the US



Army for twenty years, retiring at the rank of Lieutenant Colonel. He is a former member of the Army Acquisition Corps and a graduate of the Program Manager's Course at the Defense Systems Management College.

Professor Snider's recent publications appear in *American Review of Public Administration*, *Administration and Society*, *Administrative Theory & Praxis*, *Journal of Public Procurement*, *Acquisition Review Quarterly*, and *Project Management Journal*.

Dr. Snider received the **2009 Richard W. Hamming Annual Faculty Award for Achievement in Interdisciplinary Activities**. The selection is based on his work in leading and administering the Naval Postgraduate School's Acquisition Research Program.

Karey L. Shaffer—Program Manager, General Dynamics Information Technology, supporting the Acquisition Research Program at the Graduate School of Business & Public Policy, Naval Postgraduate School. As PM, Shaffer is responsible for operations and publications in conjunction with the Acquisition Chair and the Principal Investigator. She has also catalyzed, organized and managed the Acquisition Research Symposiums hosted by NPS.

Shaffer served as an independent Project Manager and Marketing Consultant on various projects. Her experiences as such were focused on creating marketing materials, initiating web development, assembling technical teams, managing project lifecycles, processes and cost-savings strategies. As a Resource Specialist at Watson Wyatt Worldwide in Minneapolis, Shaffer developed and implemented template plans to address continuity and functionality in corporate documents; in this same position, she introduced process improvements to increase efficiency in presentation and proposal production in order to reduce the instances of corruption and loss of vital technical information.

Shaffer has also served as the Project Manager for Imagicast, Inc., and as the Operations Manager for the Montana World Trade Center. At Imagicast, she was asked to take over the project management of four failing pilots for Levi Strauss in the San Francisco office. Within four months, the pilots were released; the project lifecycle was shortened; and the production process was refined. In this latter capacity at the MWTC, Shaffer developed operating procedures, policies and processes in compliance with state and federal grant law. Concurrently, she managed \$1.25 million in federal appropriations, developed budgeting systems and helped secure a \$400,000 federal technology grant. As the Operations Manager, she also launched the MWTC's Conference site, managed various marketing conferences, and taught student practicum programs and seminars.

Shaffer holds an MBA from San Francisco State University and earned her BA in Business Administration (focus on International Business, Marketing and Management) from the University of Montana.

A special thanks to our editors, Adrienne Malan, Jessica Moon, Steve Williams, and Lyndsee Cordes, for all that they have done to make this publication a success, to Shellee Dooley and Tera Yoder for production support, and to the staff at the Graduate School of Business & Public Policy for their administrative support. Our program success is directly related to the combined efforts of many.



Announcement and Call for Symposium Proposals

The Graduate School of Business & Public Policy at the Naval Postgraduate School announces the **8th Annual Acquisition Research Symposium** to be held **May 11-12, 2011 in Monterey, California**.

This symposium serves as a forum for the presentation of acquisition research and the exchange of ideas among scholars and practitioners of public-sector acquisition. We seek a diverse audience of influential attendees from academe, government, and industry who are well placed to shape and promote future research in acquisition.

The Symposium Program Committee solicits proposals for panels and/or papers from academicians, practitioners, students and others with interests in the study of acquisition. The following list of topics is provided to indicate the range of potential research areas of interest for this symposium: **acquisition and procurement policy, supply chain management, public budgeting and finance, cost management, project management, logistics management, engineering management, outsourcing, performance measurement, and organization studies**.

Proposals must be submitted by **November 5, 2010**. The Program Committee will make notifications of accepted proposals by **December 10, 2010**. Final papers must be submitted by **April 1, 2011**.

Proposals for papers should include an abstract along with identification, affiliation, contact information and short bio for the author(s). Proposals for papers plan for a 20 minute presentation. Proposals for panels (plan for 90 minute duration) should include the same information as above as well as a description of the panel subject and format, along with participants' names, qualifications and the specific contributions each participant will make to the panel.

Submit paper and panel proposals to www.researchsymposium.org



THIS PAGE INTENTIONALLY LEFT BLANK



Call for Research: Broad Agency Announcement

GRANTS.GOV -- NPSBAA10-002

**The Acquisition Research Program
Open until 5:00 p.m. PDST 18 June 2010**

Primary objective is to attract outstanding researchers and scholars to investigate topics of interest to the defense acquisition community. The program solicits innovative proposals for defense acquisition management and policy research to be conducted during fiscal year (FY) 2011 (1 Oct 2010 - 30 Sep 2011).

Defense acquisition management and policy research refers to investigations in all disciplines, fields, and domains that (1) are involved in the acquisition of products and/or services for national defense, or (2) could potentially be brought to bear to improve defense acquisition. It includes but is not limited to economics, finance, financial management, information systems, organization theory, operations management, human resources management, and marketing, as well as the “traditional” acquisition areas such as contracting, program/project management, logistics, and systems engineering management.

This program is targeted in particular to U.S. universities (including U.S. government schools of higher education) or other research institutions outside the Department of Defense.

The Government anticipates making multiple awards up to \$120,000 each for a basic research period of twelve months. NPS plans to complete proposal evaluations and notify awardees in mid-August 2010. The actual date of grant award will depend on availability of funds and the capabilities of the grants office. Prior year awards occurred in the August – January timeframe. Awardees may request approval of pre-award costs (up to three months), or they may request adjustments in the grant period of performance.

Full Text can be found at www.grants.gov

To locate the call quickly:

- 1) **Go to www.grants.gov**
- 2) Use **Quick Links** on the far right hand corner under **FOR APPLICANTS, Grant Search.**
- 3) Type in **NPSBAA10-002** under **Search by Funding Opportunity Number.**



THIS PAGE INTENTIONALLY LEFT BLANK





PROCEEDINGS

OF THE
SEVENTH ANNUAL ACQUISITION
RESEARCH SYMPOSIUM
THURSDAY SESSIONS
VOLUME II

**Acquisition Research:
Creating Synergy for Informed Change
May 12 - 13, 2010**

Published: 30 April 2010

Approved for public release, distribution unlimited.

Prepared for: Naval Postgraduate School, Monterey, California 93943



THIS PAGE INTENTIONALLY LEFT BLANK



Table of Contents

Keynote: Lieutenant General Thomas J. Owen, US Air Force, Commander, Aeronautical Systems Center	xv
Panel #14 – Economic Impacts of the Defense Industrial Base	402
An Analysis of the Impact of Consolidation Activity in the US and European Defense Industrial Bases	404
An Economic Analysis of Investment in the United States Shipbuilding Industry	405
Panel #15 – Applying Contemporary Economic Theory to Defense Acquisition Decision Making	428
A Three Stage Multi-attribute Procurement Auction: A Proposal for Department of Defense (DoD) Vendor Selection Decisions	429
When More is Better: Design Principles for Prediction Markets in Defense Acquisition Cost Forecasting	443
Innovations in Defense Acquisition: Implementing Information Aggregation Markets.....	465
Panel #16 – Considerations for Managing Service Oriented Architecture Projects.....	cdlxvii
An Exploration of the AT&L AV SOA System.....	469
Meaningful Cost-Benefit Analysis for Service-Oriented Architecture Projects	471
Service-Oriented Architectures and Project Optimization for a Special Cost Management Problem	476
Panel #17 – Strengthening the Industrial Base	486
Quadrennial Defense Review Priorities and Implications for Acquisition.....	487
Requisites for a Strong Defense Industrial Base	488
The Defense Budget and Defense Industry Finance	489
Panel #18 – Globalization: Cooperation and Competition in the Defense Industry	497
U.S. Export Controls and Technology Transfer Requirements—A UK Perspective	499
Global Cooperation and Competition in the Defense and Aerospace Industries	521
Panel #19 – System-of-Systems Design & Development	524
Acquisition Risks in a World of Joint Capabilities.....	525
A Technique for Evaluating Complex System of Systems Designs.....	536
System Development and Risk Propagation in Systems-of-Systems	555



Panel #20 – Managing Services Procurement in DoD	574
The Changing Face of Procurement Policy - An Innovative Approach to Competing Requirements	575
Services Supply Chain in the Department of Defense: Comparison of Acquisition Management in the Army, Navy, and Air Force	588
Panel #21 – Innovations in DoD Acquisition Policy.....	626
Improving Defense Acquisition Processes with Evidence-based Analysis: An Illustrative Case Using the DoD SBIR Program	627
Improving Defense Acquisition Decision-Making	640
US Space Acquisition Policy: A Decline in Leadership.....	657
Panel #22 – System Engineering for Project Success	668
Systems Engineering Applied Leading Indicators - Enabling Assessment of Acquisition Technical Performance.....	669
Funding for Life: When to Spend the Acquisition Pot.....	684
System Capability Satisficing in Defense Acquisition via Component Importance Measures	697
Panel #23 – Using Knowledge Value Analysis and Real Options Analysis to Improve DoD Acquisition Planning and Management.....	714
Integrating System Dynamics Modeling and Knowledge Value Added for Improved Analysis of Alternatives: A Proof of Concept Study.....	715
PEO-IWS ACB Insertion Portfolio Optimization	738
Panel #24 – Considerations for Developing and Nurturing the Future Acquisition Workforce.....	761
Determining the Optimal Staffing Level for the Acquisition Workforce: Different Models in Different Services Yield Different Results	763
Compensation, Culture and Contracts: The realities of the DoD’s Blended Workforce	765
Industry Perceptions of Department of Defense Program Manager	785
Panel #25 – Acquisition Management in a System-of-Systems World	803
Contracting for Complex Products.....	804
Research on Systems-of-Systems Acquisition.....	821
Creation of a System of Systems Portfolio Management and Technology Selection Methodology	839



Keynote: Lieutenant General Thomas J. Owen, US Air Force, Commander, Aeronautical Systems Center



Lt. Gen. Thomas J. Owen is Commander, Aeronautical Systems Center, Wright-Patterson Air Force Base, Ohio. As ASC Commander and Program Executive Officer for aircraft procurement and modernization, he leads the Air Force's Center of Excellence for Development and Acquisition of Aeronautical Systems. The center is responsible for the management of more than 556 Air Force, joint and international programs, executes an annual budget of \$23 billion, and employs a work force of approximately 10,000 people located at Wright-Patterson AFB and 38 other locations worldwide.

General Owen entered the Air Force in 1978 as a graduate of the US Air Force Academy. Early in his career he worked on B-52 Stratofortress, KC-135 Stratotanker and F-15 Eagle aircraft. The general has commanded an aircraft generation squadron maintaining F-16 Fighting Falcon and OA-10 Thunderbolt II aircraft; a maintenance squadron performing intermediate level aircraft and munitions maintenance; and a combat logistics support squadron providing F-16 and F-4 Phantom aircraft battle damage repair, supply and transportation support. He's been assigned to Headquarters US Air Force in the Pentagon two times.

General Owen was the first logistics group commander for the first and only wing operating E-8C Joint STARS aircraft. For more than three years, he was Director of the C-5 System Program Office at San Antonio and Warner Robins Air Logistics Centers. He has also served as Director for the C-17 and F-22A System Program Offices at the Aeronautical Systems Center, and Director of Logistics, Installations and Mission Support at Headquarters Air Education and Training Command. He was Commander of the Warner Robins ALC. Prior to assuming his current duties, he was Director of Logistics and Sustainment, Headquarters Air Force Materiel Command. General Owen holds Department of Defense Acquisition Corps Level III certifications in Program Management and Life Cycle Logistics.

General Owen also holds an MA in national security strategy from National Defense University and an MA in political science and international relations from Troy State University. He received his Bachelor's degree from US Air Force Academy in organizational behavior.



THIS PAGE INTENTIONALLY LEFT BLANK



Panel #14 – Economic Impacts of the Defense Industrial Base

Thursday, May 13, 2010

**9:30 a.m. –
11:00 a.m.**

Chair: James E. Thomsen, Principal Civilian Deputy, Assistant Secretary of the Navy (Research, Development & Acquisition)

Discussant: Rear Admiral David H. Lewis, US Navy, Vice Commander, Naval Sea Systems Command

An Analysis of the Impact of Consolidation Activity in the US and European Defense Industrial Bases

Nayantara Hensel, Office of the Secretary of the Navy

An Economic Analysis of Investment in the United States Shipbuilding Industry

LT Nicholas A. Meyers, US Navy, Naval Postgraduate School

Chair: James E. Thomsen, Principal Civilian Deputy, Assistant Secretary of the Navy (Research, Development & Acquisition)

Mr. Thomsen is currently the Principal Civilian Deputy Assistant Secretary (RD&A). His responsibilities include leadership of the acquisition workforce and systems engineering.

Previously, Mr. Thomsen served as the Program Executive Officer for Littoral and Mine Warfare. As PEO LMW, he had life cycle responsibility to design, produce, field, and support war fighting capability for the littoral battle space and for the global war on terrorism. Mr. Thomsen led seven program offices that comprised 224 programs ranging from ACAT I through ACAT IV and included several developmental programs that addressed urgent war fighting needs for OIF and OEF. In 2003, Mr. Thomsen was selected as the Executive Director for the Program Executive Office, Littoral and Mine Warfare where he executed the Navy's material acquisition programs for Integrated Undersea Surveillance, Naval EOD/JCREW, Naval Special Warfare, Mine Warfare Surface and Aviation, Unmanned Maritime Vehicles, Naval Anti-Terrorism/Force Protection Ashore and Afloat, and LCS Mission Modules for ASW, Mine Warfare, and ASUW.

Prior to this position, Mr. Thomsen was assigned as Head of the NAVSEA Dahlgren Division's Weapons Systems Department, directing over 550 scientists, engineers, technicians, and advancing key technical achievements in Naval Surface Weapons systems.

Mr. Thomsen was selected as a member of the Senior Executive Service (SES) in November 1998 and then named as Head, Coastal Warfare Systems Department directing all of the Littoral Warfare RDT&E programs at Naval Surface Warfare Center (Panama City) which included 360 scientists, engineers, technicians, and military personnel.



Prior to November 1998, Mr. Thomsen served as Program Manager for Mine Warfare programs, for which he was awarded the NDIA Bronze Medal for his achievements in Mine Warfare; Senior Systems Engineer for the Shallow Water Mine Countermeasures program; Project Manager for the ACAT 1D Joint US/UK Surface Ship Torpedo Defense (SSTD) program for which he received the Commanding Officer/Technical Director Award for special achievement in technical management; and Head, Torpedo Defense Systems Development Branch and served as the System Integration Agent in Submarine Torpedo Defense Countermeasure programs for PMS 415.

In the early years of his 27 year career, he held engineering positions including design engineer, test engineer, project engineer, and systems engineer for several undersea warfare programs at Carderock, Panama City, and NAVSEA Headquarters.

Mr. Thomsen received his Bachelor's degree in Ocean Engineering from Florida Atlantic University in 1981 and his Master of Science degree from Florida State University in 1986.

Discussant: Rear Admiral David H. Lewis, US Navy, Vice Commander, Naval Sea Systems Command

Rear Admiral Lewis is currently assigned as vice commander, Naval Sea Systems Command.

Born at Misawa Air Force Base, Japan, Lewis was commissioned in 1979 through the Navy ROTC Program at the University of Nebraska, Lincoln with a Bachelor of Science in Computer Science.

At sea, he served aboard USS Spruance (DD 963) as communications officer, where he earned his Surface Warfare qualification, USS Biddle (CG 34) as fire control officer and missile battery officer, and USS Ticonderoga (CG 47) as combat systems officer.

Lewis' shore assignments include assistant chief of staff for Maintenance and Engineering, commander, Naval Surface Forces, where he also served as a Charter member of the Surface Warfare Enterprise. Other ship maintenance and acquisition assignments ashore include the Navy Secretariat staff, commander, Naval Sea Systems Command staff, Aegis Shipbuilding Program Office, supervisor of Shipbuilding, Bath, and Readiness Support Group, San Diego. His major command assignment was Aegis Shipbuilding Program Manager in the Program Executive Office Ships, where he helped deliver seven DDG 51 class ships and procured another 10 ships.

Upon selection to flag rank, Lewis was serving as executive assistant to the assistant secretary of the Navy for Research, Development and Acquisition.

Lewis earned a Master of Science degree in Computer Science from the Naval Postgraduate School. He completed the Seminar Course at the Naval War College Command and Staff School, and received his Joint Professional Military Education certification. He is a member of the Acquisition Professional Community with Level III certifications in Program Management and Production Quality Management and has completed his civilian Project Management Professional certification.

His personal awards include the Legion of Merit, Meritorious Service Medal, Navy and Marine Corps Commendation, Navy and Marine Corps Achievement Medal, National Defense Service Medal, Global War on Terrorism Service Medal and various unit awards.



An Analysis of the Impact of Consolidation Activity in the US and European Defense Industrial Bases

Nayantara Hensel—Dr. Nayantara Hensel is the Chief Economist for the Department of the Navy. She provides economic guidance on growth projections, the federal budget, interest rates, unemployment, exchange rates, inflation, and the financial health of defense contractors, as well as trends in the broader economy and in the defense sector. Dr. Hensel received her BA, MA, and PhD from Harvard University, where she graduated magna cum laude and Phi Beta Kappa and specialized in finance and economics. She has taught at the US Naval Postgraduate School's Graduate School of Business and Public Policy, Harvard University, and the Stern School of Business at New York University (NYU). In the private sector, Dr. Hensel previously served as Senior Manager and Chief Economist for Ernst & Young's litigation advisory group, managing economist for the New York City office of the Law and Economics Consulting Group (LECG), and as an economist in the economic consulting arm of Marsh & McLennan. Dr. Hensel has written over 30 articles and research reports. Her recent research has focused on globalization and the US Defense industrial base (the USAF tanker competition), the role of defense mergers in improving weapons systems' cost efficiency, efficiency in IPO auctions relative to traditional processes, the factors impacting discount rates for US Marine Corps personnel, and market structure-specific and firm-specific factors impacting economies of scale and density in European and Japanese banks. She has published in the *International Journal of Managerial Finance*, the *Review of Financial Economics*, *Business Economics*, the *European Financial Management Journal*, the *Journal of Financial Transformation*, and *Harvard Business School Working Knowledge*. She is the Chair of the Financial Roundtable for the National Association of Business Economists (NABE) and is one of 34 elected members to NBEIC, a group composed of the top corporate economists in the US. Dr. Hensel has given seminars at a number of institutions and has appeared on CNBC, Bloomberg Radio, and CNNMoney.

Executive Summary

The defense industrial base in the US has witnessed many changes over the past twenty years, following the end of the Cold War, and has been reshaped by a variety of significant forces. This project examines the impact of macroeconomic conditions (the federal budget, labor market conditions) and the recent shift in defense priorities toward irregular warfare, on the defense industrial base, as well as the globalization of the defense sector. First, the analysis examines trends in the federal debt and deficit, as well as defense spending, over the past fifty years, and assesses the possible pressure placed on future defense spending by fiscal austerity in the coming years, including the emphasis on certain areas of the defense industrial base in the recent budgets and in the Quadrennial Defense Review. Second, the project examines the impact of stagnant regional labor markets on Congressional concerns over cutbacks in particular defense programs. Third, the project assesses concerns over the reduction in skill sets in certain areas of the defense industrial base, due to possible reduced procurement rates, as well as the shift in defense priorities toward irregular warfare. The analysis examines several areas, including the solid rocket motors defense sub-sector, as well as the impact of shifts in recent USN thirty-year shipbuilding plans on the shipbuilding industrial base. Fourth, the analysis examines areas of growth in the defense industrial base due to the shift in priorities toward irregular warfare. Fifth, the project assesses the involvement of foreign defense contractors in the US defense industrial base and the success of global supply-chain alliances between foreign and US defense firms. Finally, the project analyzes the continued demand from foreign markets for conventional weapons systems and the growth in defense industrial bases in other countries.



An Economic Analysis of Investment in the United States Shipbuilding Industry

LT Nicholas A. Meyers, US Navy—Lieutenant Nicholas Meyers is an MBA student in the Graduate School of Business & Public Policy at the Naval Postgraduate School. LT Meyers is specializing in Financial Management and will graduate in June 2010. He received a Bachelor of Science in Computer Engineering, cum laude, from Virginia Tech in 2002. An active-duty Submarine Officer, LT Meyers served his Junior Officer tour aboard USS Columbia (SSN-771) in Pearl Harbor, HA, and also served as a Submarine Watch Officer and Regional Employment Officer at Submarine Forces Headquarters (COMSUBFOR) in Norfolk, VA. LT Meyers and his wife Sarah reside in Monterey with two children—daughter Maren Beth, and 3-month-old son Benjamin Aaron.

LT Nicholas Meyers, USN
1298 Spruance Road
Monterey, CA 93940
831.333.0317 (home)
808.294.4903 (cell)
nmeyers@nps.edu

Abstract

Amidst the global economic recession of recent years and sizeable injections of federal stimulus packages, the Navy's budget for ship construction (SCN) has experienced only modest real growth. While both the 2010 Quadrennial Defense Review and the 30-year shipbuilding plan have reaffirmed a fleet size goal of 313 ships, some suggest that \$20 billion or more per year is needed to attain these fleet numbers. This research has analyzed the United States' shipbuilding industry as a potentially rewarding source of economic stimulus and benefit through similar or identical measures used by economists at Oxford Economics. First, direct and indirect (supply-chain) monetary impacts from the "shipbuilding and repair" sector were analyzed using US Bureau of Economic Analysis input/output data and a Carnegie-Mellon University model of a Leontief inversion process. This sector was then compared with five alternative investments. Second, the direct and indirect benefits of the shipyard-related labor market was analyzed using data collected from the Bureau of Economic Analysis, Naval Sea Systems Command, and the American Shipbuilding Association. Finally, measures of Capital Intensity and Capacity were applied to the financial statements of the nation's largest private owners of shipyards—General Dynamics and Northrop Grumman. The results suggest that US shipbuilding generates benefits comparable to alternative investments, while supporting more labor, and highly skilled jobs, than alternatives. In addition, high levels of capital intensity in shipbuilding suggest that a decline in demand may yield a permanent loss of US productive capacity. Finally, excess capacity throughout the industry shows a clear ability to absorb an increase in demand, providing prompt and immediate impact on sustained economic recovery.

Introduction

Purpose of Study

In 2008 and 2009, the United States' economy has struggled with what has widely been described as "the worst economic crisis since the Great Depression." Specifically, the national economic data shows a reduction of 1.7 % in real gross domestic product (GDP),



measured in constant 2005 dollars, since the beginning of calendar year 2008 (Bureau of Economic Analysis, 2009). Although this contraction may seem slight, this is the first six-quarter period since 1982 that the national GDP growth has been negative. Now, with the nationwide unemployment rate near 10%, the United States has lost over 7.3 million jobs since the start of the damaging recession (Homan, 2009). Some economists are predicting recovery in 2010, but national leaders and decision-makers continue to look to federal government spending as a means of stimulating job growth, injecting stability, and sustaining recovery.

In a series of efforts to mitigate drastic economic decline, the US Congress passed a \$700 billion *Troubled Asset Relief Program* (TARP) package in October of 2008, followed by a \$787 billion *American Recovery and Reinvestment Act* on February 13th, 2009 (Recovery.gov, 2009). Of the initial TARP package, about \$550 billion has been committed to various financial firms, banks, and institutions throughout the country; so far, \$70.1 billion has been returned to the Treasury (Ericson, He & Schoenfield, 2009). The Federal Reserve and the White House continue now to seek proper locations for depositing large sums of federal dollars as a means of ensuring continued consistent recovery of our national economic forecasts. Just as recently as December of 2009, the *New York Times* featured a front-page article in which White House economist Jared Bernstein suggests that the administration is considering an additional \$150 billion in stimulus spending, of which \$50 billion could be invested “in infrastructure projects alone such as roads, bridges, and water projects” (Pace, Taylor & Elliott, 2009). Clearly, national leaders are convinced that boosting federal spending is one of the best tools for ensuring that America's \$14.3 trillion¹ economy remains healthy and growing at a stable, sustainable pace. The questions now being discussed in various offices and conference rooms throughout Washington, DC, and the country as a whole include robust debates about where to invest these funds. What effects would a \$1 trillion health care package have on our weakened economy? Where are the benefits of the *American Recovery and Reinvestment Act*; where have they manifested? The executive branch claims to track every dollar spent under this \$787 billion umbrella, but how is that spending really benefiting the economy? Other, perhaps equally important, questions exist for industries and sectors yet to benefit from federal spending packages and stimulus measures—what *could* investments in those sectors be doing to improve the economy?

One important industry that has not received direct funding from government intervention in the current recession has been the US shipbuilding sector. A website search for “shipbuilding” at the federal government’s website designed to provide transparency to American citizens reveals that a mere \$132,000 of the \$787 billion package has been allocated to a company called Horizon Shipbuilding in Alabama (Recovery.gov, 2009).² This \$132,000 payment from the Department of Transportation is the only search result; clearly, the outcome demonstrates that not even one tenth of one percent of the *American Recovery and Reinvestment Act* has been invested in shipbuilding companies.

Those with an interest in the US shipbuilding industry believe that their particular sector of manufacturing has a significant ability to provide economic stimulus, or substantial

¹ Based on gross domestic product (GDP), seasonally adjusted, annualized amount for 3rd quarter 2009, from the Bureau of Economic Analysis (www.bea.gov)

² Search was conducted at www.recovery.gov, whose stated mission statement is to “provide easy access to data related to *Recovery Act* spending and allow(s) for the reporting of potential waste, fraud, and abuse.”



return-on-investment for national decision-makers and taxpayers alike. The purpose of this thesis study is to determine, *what is the return or benefit to the US national economy for federal expenditures in the shipbuilding sector?* Common economic models, discussed in detail in the pages that follow, will be applied to pertinent sector data in order to answer this important question. As politicians seek to stimulate and sustain US economic growth, they hope to create or maintain jobs, expand national gross domestic product, and provide a lasting resource for future economic potential; investments in shipbuilding ardently accomplishes all three goals, as this study will seek to demonstrate.

Problem

Problem 1: US Navy Fleet Size: An Uncharted Goal

The Chief of Naval Operations (CNO) has repeatedly affirmed a commitment to a United States' naval fleet of at least 313 warships (McIntire, 2009). However, in the FY 2011 30-year shipbuilding plan, the US naval fleet does not include 313 ships until the year 2020 (Director, Warfare Integration (OPNAV N8F), 2010). One profound problem is the sharp discrepancy between these CNO estimates of national needs for our naval fleet, versus the projected fleet decline if funding for ship construction remains constant in real dollars. An estimate from the American Shipbuilding Association suggests that our fleet could reach a mere 180 ships total if additional funding for ship construction is not received.³ Another study by the Center for Strategic and Budgetary Assessments (CSBA), suggests that the Congressional appropriation for "Shipbuilding and Conversion, Navy (SCN)" would have to be funded to levels of about \$20.4 billion in order to achieve the Navy's desired force structure in future years (Work, 2009). Moreover, the same CSBA study references recent Congressional Budget Office estimates that a total of \$22.4 billion per year would be required to reach a fleet size of 313 ships. In stark contrast to these projected levels for achievement of the CNO's fleet size goal, \$12.7 billion was the total funding of the fiscal year 2009 SCN account; based on CSBA estimates, the effort was underfunded by about 35% (Department of the Navy, 2008). Additionally, throughout 2040, "to be consistent with expected future defense budgets, the Department of the Navy's annual shipbuilding construction (SCN) budget must average no more than \$15.9 billion per year in FY2010 dollars" (Director, Warfare Integration (OPNAV N8F), 2010).

³ The prediction is captured in the "unilateral Navy trends" graph, acquired from the American Shipbuilding Association, December 2009. Included on next page.



Unilateral Navy Disarmament

12 Ships A Year Will Meet the Requirement

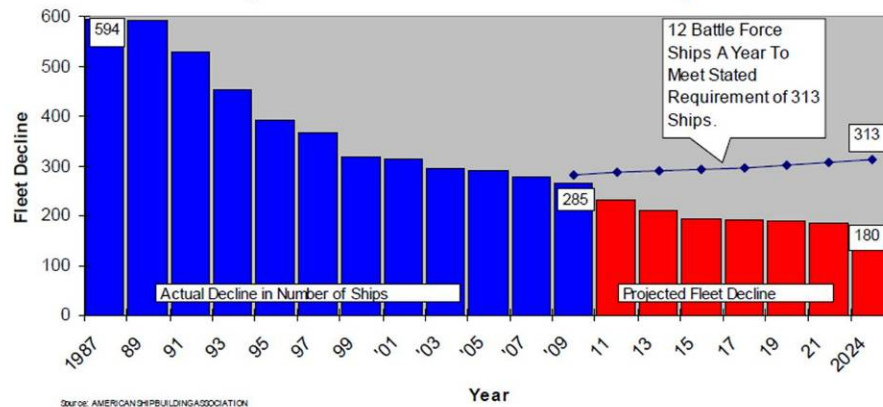


Figure 1. Unilateral Navy Disarmament

In summary of the first disconnect between projected fleet needs and recent funding of ship construction, the ostensible likelihood is that, barring any geopolitical events that may punctuate the national security equilibrium, the United States Navy is not likely to reach a fleet of 313 ships before the year 2020. In addition, the academic and practical arenas of shipbuilding cost growth and projected fleet size funding estimates have been thoroughly explored by talented minds with reliable experience. For these reasons primarily, this thesis study will not explore issues of (1) appropriate fleet size to meet national security needs, or (2) the rising costs of ship construction as it impacts efforts to reach projected needs. Although both of these issues are considered important, thorough and credible studies by congressional experts such as Mr. Ron O'Rourke⁴, RAND, and others have been and currently are being conducted within these arenas; this thesis will focus elsewhere.

Problem 2: What Does Shipbuilding Do for the Economy? How?

A second important problem exists within these topics of US naval fleet size, national economic woes, and ship construction—one more obscure, and perhaps with a solution more potentially rewarding. The question referred to is this: what are the economic benefits of building ships? When comparing alternative options for investment of federal taxpayer dollars, building highways versus bailing out banks for example, several important questions appear at the forefront of consideration. First, what is the health and capacity of the industry being considered for receipt of billions of dollars? Could the sector accept the billions of dollars of additional funding and apply them in some meaningful manner to an economic benefit of others? Are there vendors or resources whose financial health depends upon the sector considered? For instance, building highways requires not only large pools of available labor, but also purchases from blacktop/concrete producers, perhaps equipment rentals for steamrollers and forklifts, and other suppliers who would benefit from increased demand of their products. Are there similar supply-chain benefits for the builders of ships?

⁴ Mr Ronald O'Rourke is a "specialist in naval affairs" with the Congressional Research Service, and has published numerous studies regarding the rising costs of Navy ships for the US Navy, as well as analysis of other nation's ships procurement programs; most are available at www.crs.gov.

Investments of federal tax dollars should be allocated to the sectors where it would economically benefit the most people, or maximize the return-on-investment for the federal government.

When evaluating a particular industry for its ability to benefit national economic recovery and growth, one should consider the channels through which the benefits manifest. In determining channels of impact, quantitative multipliers may be calculated. Once a trusted and scrutinized economic multiplier is available for each investment option, public sector decision-makers could be as well informed as private sector investment bankers or venture capitalists, who seek to deposit their wealth where it will multiply the greatest, earn the most rewards, and pay dividends to their stakeholders for future periods. The national politicians seeking stabilization and growth of the US economy could evaluate their problem of economic return for various courses of investment using a similar framework; their stakeholders are all US citizens, and their wealth is the measure of national GDP.

Organization of Thesis Study

What's Not Included, and Why?

In both quantitative and qualitative analysis, this study will consider only the “US shipbuilding industry” investment option available to policymakers. One aspect of the study's contribution to national information regarding economic investment in shipbuilding will be found in its analysis of the *lifecycle benefit* of a vessel's national economic impact. Much detailed research has been conducted already in an effort to capture and quantify *lifecycle costs* of ship procurement programs. Thus, issues of rising construction and procurement costs of Navy ships will not be a focus of this study. Rather, *lifecycle benefit to the regional and national economy* will be explored, quantified where possible.

Both the 2010 *Quadrennial Defense Review (QDR)* and the Navy's 30-year shipbuilding plan were completed in February of 2010. The mix of vessels to be procured has been planned, and evaluation of shipbuilding costs is being thoroughly scrutinized by national authorities on the subject such as Ronald O'Rourke, with the *Congressional Research Service*. With the *QDR* and the 30-year shipbuilding plan both published along with the President's Fiscal Year 2011 Budget in February, much light has been shed on the path ahead for US shipbuilding and naval fleet size. This thesis study will not attempt to make a contribution to the analysis of rising ship costs, proper mix and type of vessels to be procured, or consideration of the national funds available for ship construction. Rather, the study will focus on (1) in what ways, and (2) how the shipbuilding industry benefits regional economies and the national economic health.

Methodology

General Approach

In a September 2009 study entitled *The Economic Case for Investment in the UK Defense Industry*, researchers and economists at Oxford Economics in London developed a detailed framework for analyzing the economic contribution of various industries (Oxford Economics, 2009). This study will use much of the Oxford study's framework as it analyzes the economic returns of the “shipbuilding and repair” sector through at least four lenses:

Monetary impact – using input/output analysis to analyze the direct and indirect channels of the shipbuilding's sector on the US economy



Labor market impact – jobs supported by US shipbuilding, and the relative skill levels of those jobs; the regional distribution of those jobs throughout the country

Capital Intensity – “sectors that invest the most in capital and labor present the largest potential for losses if they fail” (Oxford Economics, 2009).

Capacity measures & a rapid return? – “in order for an increase in Government procurement to have an immediate impact on the economy a sector must have sufficient spare capacity to absorb the additional demand” (Oxford Economics, 2009).

In order to best understand the methodology used by Oxford Economics, meetings were held at the London Office in December of 2009, with the two economists who were mainly responsible for the study’s content—Mr. Andrew Tessler and Mr. Pete Collings. Additionally, interviews were conducted with US acquisition professors, shipbuilding industry leaders, and several distinguished economists both within and outside of the US Department of Defense.

“Free Market” Concerns

In several months of research and interviews with various economists, some have expressed fundamental ideological concerns regarding the use of government spending as a means of stimulating the economy. The idea that government spending creates a multiplier effect for economy benefit was based on the economic theory of John Maynard Keynes, and published by Richard Kahn in 1931 (Kahn, 1931). Much academic debate and theory continues to permeate today’s economics. In today’s environment, a prominent professor of economics at Harvard University, Mr Robert Barro, has conducted research demonstrating that there is “no evidence of a Keynesian multiplier effect” for stimulus spending, and has published his view that “defense-spending multipliers exceeding one likely apply only at very high unemployment rates, and nondefense multipliers are probably smaller” (Redlick, 2009). Still other prominent economists disagree about the beneficial effects of Keynesian spending. The *Journal of Post Keynesian Economics* exists where scholars can publish their research and works based on the theories of Keynesian multipliers and Keynes’ ideas of stimulus.

One distinguished economist with Stanford University’s Hoover Institute and others with the Naval Postgraduate School have suggested that the United States Government could acquire ships more efficiently (at a lower cost) by allowing them to be produced overseas, where there may be a comparative advantage for ship construction. Although perhaps economically sound, national decision-makers widely agree that US national security requires maintaining the ability to build warships on American soil. Once agreed that the capability to produce US warships on American soil is vital to US national security interests, the benefits, or economic returns of doing so ought to be well known. In the *Journal of Post Keynesian Economics*, findings published in 2005/2006 reveal that “a rise in defense spending had a favorable impact on GDP and employment, but led to larger trade and budget deficits” (Atesoglu, 2005-6). Although there is much political and economic debate on the merits of government spending and investment, this research will build upon the work of credible, established Nobel laureate economists⁵, and closely follow the methodology employed by researchers at Oxford Economics.

⁵ Wassily Leontief received “The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 1973,” Paul Krugman, who widely agrees with Keynes’ theories, received the prize in 2008.



Monetary Impact—Multipliers: Direct, Indirect, Induced

Input/Output Analysis

Input-output economic analysis is a Nobel Prize-winning analytical framework developed by Professor Wassily Leontief in the late 1930s (Miller & Blair, 1985). All economic activity within a country is divided into sectors or industries. In the United States, those sectors are identified using the North American Industrial Classification System (NAICS) codes.⁶ Inter-industry transactions are then measured for a specific time period (one year) in constant monetary terms (the US dollar). The results, known as benchmark data, are represented in a matrix consisting of outputs listed in rows, and inputs listed in columns. The format allows analysis of how one industry's outputs are dependent upon inputs from all other sectors of the economy. The United States' Bureau of Economic Analysis last collected such economy-wide benchmark data for the US economy in 2002; a revised version was last published in April 2008 (Bureau of Economic Analysis, 2008).

Once in possession of benchmark economic data for the economy as a whole, a series of specific steps may be performed in order to identify a specific sector's impact on the economy. First, the flow from sector i to sector j is defined as z_{ij} . Next, the variable X_j is chosen as the total gross output of the individual sector j in the given year. From these variables, a technical coefficient, a_{ij} is calculated as:

$$a_{ij} = \frac{z_{ij}}{X_j}$$

The resulting coefficient then represents the dollar value of inputs from sector i required for every dollar of output from sector j . The system is designed as to provide constant returns to scale. In other words, a_{ij} is a fixed relationship; when output from sector j is doubled, it is assumed that the inputs required from sector i would also be doubled. Economies of scale in production are thus ignored; the Leontief system is strictly a linear model. Furthermore, the inter-industry flows from i to j for a given year depend entirely and exclusively on the total output of sector j for that specific year (Miller & Blair, 1985).

Rather than manually performing the matrix algebra required to analyze the impacts of a certain sector, a software model developed by researchers at Carnegie Mellon University performs a Leontief inverse on the portion of the larger matrix pertinent to the sector chosen. The model, originally created in 1995, is called Economic Input-Output Life Cycle Assessment (EIO-LCA), and is "comprised of national economic input-output models and publicly available resource use and emissions data" (Carnegie Mellon, 2008). Only the economic results will be used in this study; the environmental impact will not be considered.

With a credible reputation based on the Nobel Prize winning theory of Wassily Leontief and reliable data from the BEA, "the EIO-LCA method has been applied to economic models of the United States for several different years, as well as Canada, Germany, Spain, and select US states." The on-line tool has been accessed over 1 million times by researchers, LCA practitioners, business users, students, and others." Additionally, the input-output analysis method has been "used extensively for planning throughout the world" (Carnegie Mellon, 2008).

⁶ NAICS is the "standard used by Federal statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data related to the US business economy."

Direct & Indirect Multipliers

By considering various channels of impact, economic multipliers may be calculated for three distinct areas of the shipbuilding industry's overall economic impact: direct effects, indirect effects, and induced effects. Direct impacts are employment and activity in the sector itself—US shipbuilding. Indirect impacts are defined as “employment and activity supported down the supply chain, as a result of a sector’s companies purchasing goods and services from” suppliers (Oxford Economics, 2009). For example, when a shipyard is building a new Littoral Combat Ship (LCS), it may order a fire-control system to be installed that was designed in California. That same system may have been built with components from Washington State. The purchase of various equipment and supplies from vendors, as well as jobs and sales at those vendors’ offices may be quantified as indirect impacts for investment in the shipbuilding industry. Finally, *induced* impacts are also of economic importance to the study of ship construction. Oxford Economics defines induced impacts as “employment and activity *supported by the consumer spending* of those employed in the sector or in its supply chain.” For instance, the manufacturer of a component ordered by the shipyard for construction of a new vessel has additional revenue from the sale of that component; he spends this revenue in his local economy buying everyday goods and services, which increases benefits to local economic growth. This induced analysis considers a wide variety of industries and activities throughout the United States, and relies on creation of an economic multiplier for its quantification.

Other Sectors to be Compared

The “shipbuilding and repairing” sector will henceforth be referred to simply as the “shipbuilding” industry. Per NAICS labeling, shipbuilding is a sub-sector of the (336xxx) group labeled “vehicles and other transportation equipment.” Comparisons of Leontief model output will be analyzed and contrasted with five other sectors of the US economy:

- Automobile manufacturing (336111)
- Aircraft manufacturing (336411)
- Military Armored Vehicles and tank parts manufacturing (336992)
- Nonresidential manufacturing structures (230102)
- Health care: offices of physicians, dentists, health care practitioners (621A00).

These five sectors were chosen to include three other subcategories of manufacturing transportation vehicles, a more general manufacturing alternative, and also a service-based industry for comparison.

Estimation of Induced Multipliers

In addition to the direct and indirect economic effects to be calculated using the Carnegie Mellon model, *induced* effects should also be considered and quantified. The induced impacts of activity within a sector are “employment and activity supported by the consumer spending of those employed in the sector or in its supply chain. This helps to support jobs in [US] industries that supply these purchases and includes jobs in retail outlets, companies producing consumer goods and in a range of service industries” (Oxford Economics, 2009). Since the induced effects are the most difficult to quantify, data from previous studies of US and UK shipbuilding industries will be reviewed. Based on the recommendation of economist Andrew Tesler at Oxford Economics, the induced multiplier for US shipbuilding will be estimated as a fraction of the indirect multiplier. Based on the ratio of induced effects to indirect effects for similar studies, as well as the basic



consumption multiplier, this research will lead to an inference about a reasonable range of an induced multiplier for US shipbuilding.

Estimations of Employment Supported

Based on the work of Garnick and Drake in the 1970s, the Bureau of Economic Analysis (BEA) has published a handbook for users of its Regional Input-Output Multipliers System (LECG, LLC, 2002)⁷. The process of using the BEA's system to derive regional multipliers is summarized concisely in the 2002 LECG report for the American Shipbuilders Council:

The RIMS II method for estimating regional Input-Output multipliers can be viewed as a three-step process. In the first step, the producer portion of the national Input-Output table is made region-specific by using four-digit SIC location quotients (LQ's). The LQ's estimate the extent to which input requirements are supplied by firms within the region. RIMS II uses LQ's based on two types of data: BEA's personal income data (by place of residence) are used to calculate LQ's in the service industries; and BEA's wage-and-salary data (by place of work) are used to calculate LQ's in the nonservice industries.

In the second step, the household row and the household column from the national Input-Output table are made region-specific. The household row coefficients, which are derived from the value-added row of the national Input-Output table, are adjusted to reflect regional earnings leakages resulting from individuals working in the region but residing outside the region. The household column coefficients, which are based on the personal consumption expenditure column of the national Input-Output table, are adjusted to account for regional consumption leakages stemming from personal taxes and savings.

In the last step, the Leontief inversion approach is used to estimate multipliers. This inversion approach produces output, earnings, and employment multipliers, which can be used to trace the impacts of changes in final demand on directly and indirectly affected industries.

Rather than manually performing the matrix algebra and Leontief inversion, the results of the Carnegie-Mellon Economic Input-Output Life Cycle Assessment model will once again be utilized. The process allows for a distinct number of jobs to be calculated for each dollar amount of increased output from a sector.

Labor Market Impact

Highly Skilled Jobs

Many of the workers involved in ship construction, and modernization have been training for years to earn the specific qualifications necessary to perform those tasks. To be a nuclear plant welder in the United States, for example, "one must be cleared by the FBI, undergo drug and alcohol testing, and pass a psychological screening. These criteria are above and beyond welding certification, diving certification, and special training required of all nuclear plant personnel (Hancock, 2003)." The nuclear welders and construction personnel who build our aircraft carriers and submarines are not an immediately renewable

⁷ Regional Input-Output Multipliers System (RIMS II) is explained in Appendix C of the 2002 LECG report.

resource. In other words, if they are eliminated from the workforce due to drastic drops in demand for their services at the “big six” shipyards, then there are at least two formidable and unfavorable results. First, if the US military suddenly has an increased demand for specialized labor in nuclear or conventional ship construction (war), then we will not have that capacity available to be utilized. We may have to actually outsource those jobs to other country, which is particularly dangerous and difficult in matters of national security and weapons systems construction. Secondly, the atrophy of the workers’ skills in industry combined with the graying of the workforce may actually lead to a regression of the “knowledge economy” of this sector of the US defense and shipbuilding industries, leading to a larger-scale contraction (National Defense Research Institute (RAND), 2006). The principle of a knowledge economy is, in brevity, explanation of the use of knowledge itself as a product or tool producing an economy benefit (Drucker, 1992). For instance, the training, experience, and skill-level of an individual welder or shipyard worker has some inherent economic value, which can be quantified in calculating the sum of the industry or activity’s economic worth.

Capital Intensity and Excess Capacity—“What If?”

In researching the unique aspects of the US shipbuilding industry as it compares to other defense activities, Dr. Nayantara Hensel, a former professor at The Naval Postgraduate School and currently Chief Economist in the Assistant Secretary of the Navy (Financial Management & Comptroller Office) highlighted the high capital intensity and sunk investments of infrastructure existing within the shipbuilding sector. The facilities and infrastructure themselves become “economic waste” if the existing capacity is not utilized by providing an appropriate demand signal (Hensel, 2009). The same principle is summarized nicely in the Oxford Economics report on investment in the UK defense industry: “sectors that invest the most in capital and labour present the largest potential for losses, if they fail.” In summary, unlike shopping malls and retail centers, shipyards (as they are highly both capital and labor intensive) are unable to be readily converted to some other economic activity, if they fail. Rather, they become “waste.” There is, therefore, an inherent opportunity cost of failing to utilize the existing capacity—the current market values of the facilities and technology themselves. Acceptance of this principle that irrevocable waste results from failure to utilize sectors with high capital intensity, combined with the clear evidence of the shipbuilding industry’s investment in capital plants and equipment, supports the claim that basic funding levels to sustain the industry’s existence at current levels is economically viable and preferable (Booth, Colomb & Williams, 2008).

Using public data from the released “10k” financial statements, capital intensity will be calculated as:

$$\frac{\text{net capital stock}}{\text{revenue for 2009}}$$



This ratio provides “a measure of a firm's efficiency in deployment of its assets, computed as a ratio of the total value of assets to sales revenue generated over a given period. Capital intensity indicates how much money is invested to produce one dollar of sales revenue.” Moreover, “a decline in a capital intensive industry may mean a permanent loss of productive capacity” (Oxford Economics, 2009).

Capacity Measures and a Rapid Return?

“In order for an increase in Government procurement to have an immediate impact on the economy a sector must have sufficient spare capacity to absorb the additional demand” (Oxford Economics, 2009). Published data from the US Department of Commerce, Bureau of Economic Analysis, and corporate “10k” financial statements will be collected and analyzed in order to determine if the US shipbuilding industry could absorb an increase in demand and provide a timely return for investment.

Results

Input/Output Multiplier Analysis

Carnegie Mellon University's Economic Input-Output Life Cycle Assessment (EIO-LCA) model was used to perform a Leontief inverse solution based on 2002 US Benchmark data from the Bureau of Economic Analysis (BEA), with the following results obtained (Carnegie Mellon, 2008). First, it should be noted that the latest United States Benchmark data from the Bureau of Economic Analysis is 2002 data that was last updated in 2008. The model's software calculates a coefficients matrix based on the input-output data for the US economy. By isolating a single sector of the economy, and choosing a given level of output or production from that sector, direct and indirect economic activity estimates are generated. The “shipbuilding and repair” sector (NAICS code 336611) was selected for analysis, with a presumed increased production from that sector of \$100 million. In other words, an injection of \$100 million was entered into the model. One possible source of an additional ship production demand of \$100 million would be government orders for US Navy vessels. However, this particular model makes no distinction between military and civilian contracts, nor between Navy and commercial shipbuilding. If the market were to demand an additional \$100 million in commercial ship construction, the economic activity estimates would be the same. Since the Leontief function is a linear model, output results will vary proportionally with those generated below. Indeed, Leontief models are always linear (Leontief, 1966). For instance, entering a \$1 billion increased demand output in to the model will yield results that are ten times higher than those below, while inputting \$50 million will yield results that are half of those below.

Table 1. Total and Direct Economic Effects of \$100 Million Output from “ShipBuilding and Repairing” Sector

<u>NAICS code</u>	<u>Sector Description</u>	<u>Total Economic (\$ mil)</u>	<u>Direct Economic (\$ mil)</u>
	<i>Total for all sectors</i>	<i>209.</i>	<i>157.</i>
336611	Shipbuilding and repairing	99.9	99.9
420000	Wholesale trade	6.7	3.6
550000	Management of companies and enterprises	6.5	3.4
333618	Other engine equipment manufacturing	5.6	4.9
331110	Iron and steel mills	3.3	1.7
533000	Lessors of nonfinancial intangible assets	3.1	2.4
541610	Management consulting services	2.5	1.8
52A000	Monetary authorities and depository credit intermediation	2.4	1.2
523000	Securities, commodity contracts, investments	2.18	1.33
531000	Real estate	2.15	0.407

Direct Economic Effects

In the first row of the above Table 1, labeled “total for all sectors,” direct economic effects of \$157 million represent the dollar amounts of purchases made by the “shipbuilding and repairing” sector in order to manufacture its final product (a ship). This \$157 million includes the input value of \$100 million increased economic activity for the shipbuilding and repairing sector, which is shown (minus rounding error of 0.1) in the second row of the table. So, the sector purchases \$57 million worth of products (goods and services) from other sectors in order to make \$100 million worth of output.

The shipbuilding and repair sector ranks third of the six sectors considered, when ranked by direct economic effects, as shown in table IV-2 below:

Table 2. Direct Economic Impact of an Additional \$100 Million Output from Sector

<u>Sector #</u>	<u>Sector Description</u>	<u>Direct Economic Impact</u>
336111	automobile manufacturing	\$174
336411	aircraft manufacturing	\$165
336992	mil. Armored vehicles & tank parts manufacturing	\$160
336611	shipbuilding and repairing	\$157
230102	nonresidential manufacturing structures	\$141
621A00	offices of physicians, dentists, health care practitioners	\$135

Value Added

The difference between the \$100 million output from shipbuilding and the \$57 million of inputs it requires is the value added by the “shipbuilding and repairing sector” itself. The value added represents “compensation of employees, taxes on production and imports less subsidies, and gross operating surplus. Value added equals the difference between an industry’s gross output (\$100 million) minus the cost of its intermediate goods that are purchased (such as energy & raw materials)” (Carnegie Mellon, 2008). For instance, once the raw materials and services are purchased from other sectors, the value of the skilled labor and contribution from the shipyards themselves totals \$43 million. Stated differently, the \$43 million in value added is one (direct) component of an increase in Gross Domestic Product (GDP) as a result of the additional \$100 million of output.

Table 3 and Figure 2 below show the value added (contribution to GDP) by sector of the US economy considered, for a \$100 million increased output from that sector. The shipbuilding sector ranks fourth of the six sectors considered, when ranked by economic value added.

Table 3.

<u>Sector #</u>	<u>Sector Description</u>	<u>Increased Production Output</u>	<u>Amount of Direct Purchases</u>	<u>Value Added (difference)</u>
336111	automobile manufacturing	\$100	\$74	\$26
336411	aircraft manufacturing	\$100	\$65	\$35
336992	mil. Armored vehicles & tank parts manufacturing	\$100	60	\$40
336611	shipbuilding and repairing	\$100	\$57	\$43
230102	nonresidential manufacturing structures	\$100	\$41	\$59
621A00	offices of physicians, dentists, health care practitioners	\$100	\$35	\$65

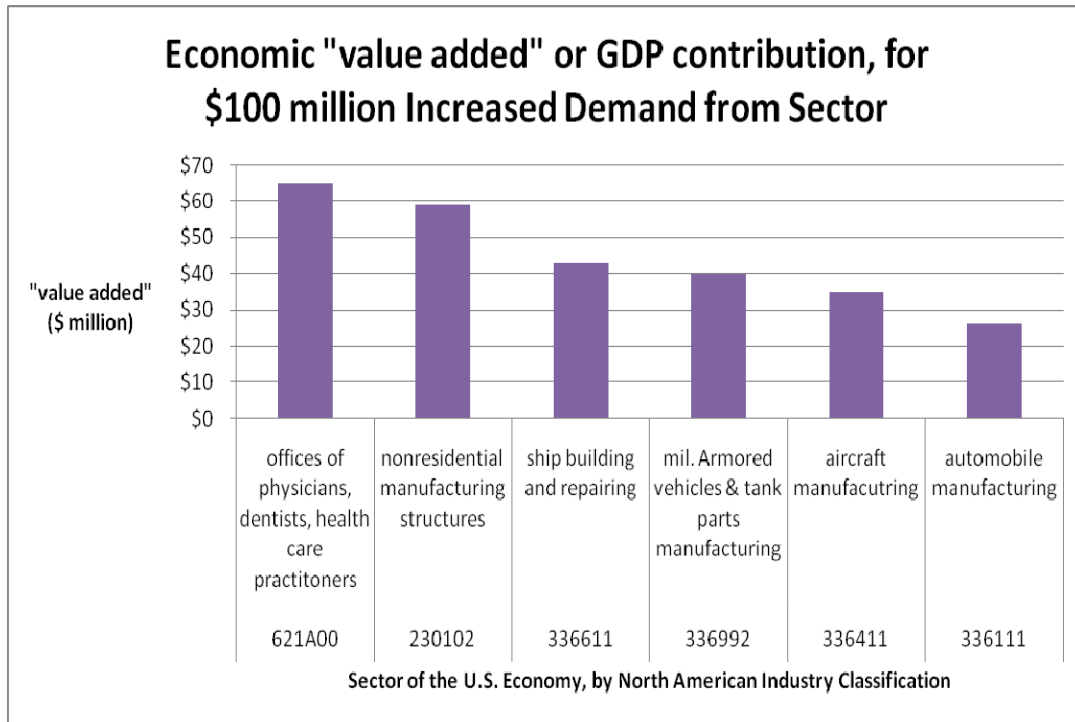


Figure 2. Economic "Value Added" or GDP Contribution, for \$100 Million Increased Demand from Sector

In these analyses, shipbuilding and repair "stands out" as a sector which is a very nearly split between the two main factors of production needed to generate additional output—materials and labor (purchases and value added). Whereas automobile and aircraft manufacturing are ranked number one and number two respectively in terms of direct economic effects, this ranking reflects a high degree of automation in their manufacturing processes. Most of the generated direct activity is due to purchases of materials these industries must make in order to manufacture their finished goods.

When the results are analyzed in terms of value added by the industry itself, the ranking of the six sectors considered is nearly inverted. In other words, "offices of physicians, dentists, and health care practitioners" which is first when ranked by value added, was last in total direct effects. Shipbuilding remains in the middle of the group when ranked by value added, since, as a sector, it requires about 57% of materials (\$53 million/\$100 million), and 43% labor as components of the additional \$100 million output. One may conclude that shipbuilding represents a "healthy balance" between these two contributing factors of production, providing stimulation of the economy through both purchases and wages.

Total Economic Effects

For "shipbuilding and repair" a total economic impact of \$209 million, as presented in the first row of Table 1 above, represents the total purchases by all other sectors of the economy resulting from an additional \$100 million output from shipbuilding and repairing. The \$209 million includes the direct purchases made by the shipbuilding and repair sector itself, and also the indirect purchases further up the supply chain; the materials and services needed to produce the goods sold to the shipbuilding and repair sector. Included

within the \$209 million of activity is the \$100 million of increased final output from shipbuilding. Figure 3 below shows how the total \$209 million is divided.

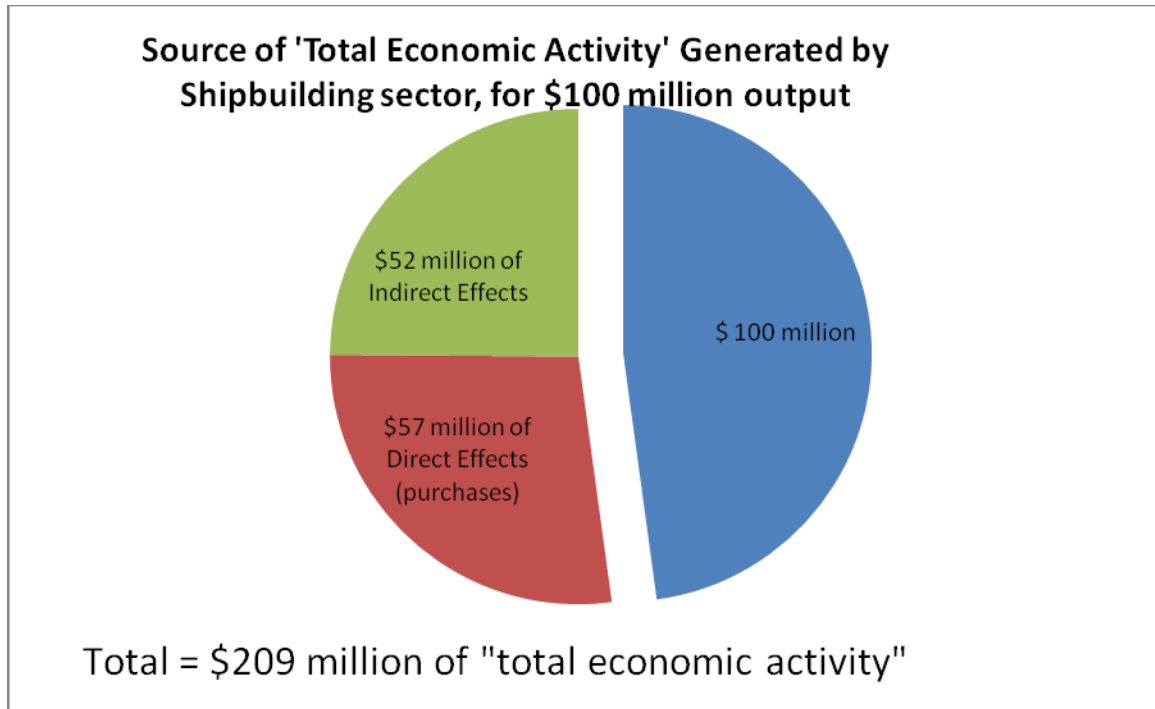


Figure 3. Graphical Representation of the Total and Direct Economic Effects Generated by an Additional \$100 Million Output from Six Different Sectors of the US Economy

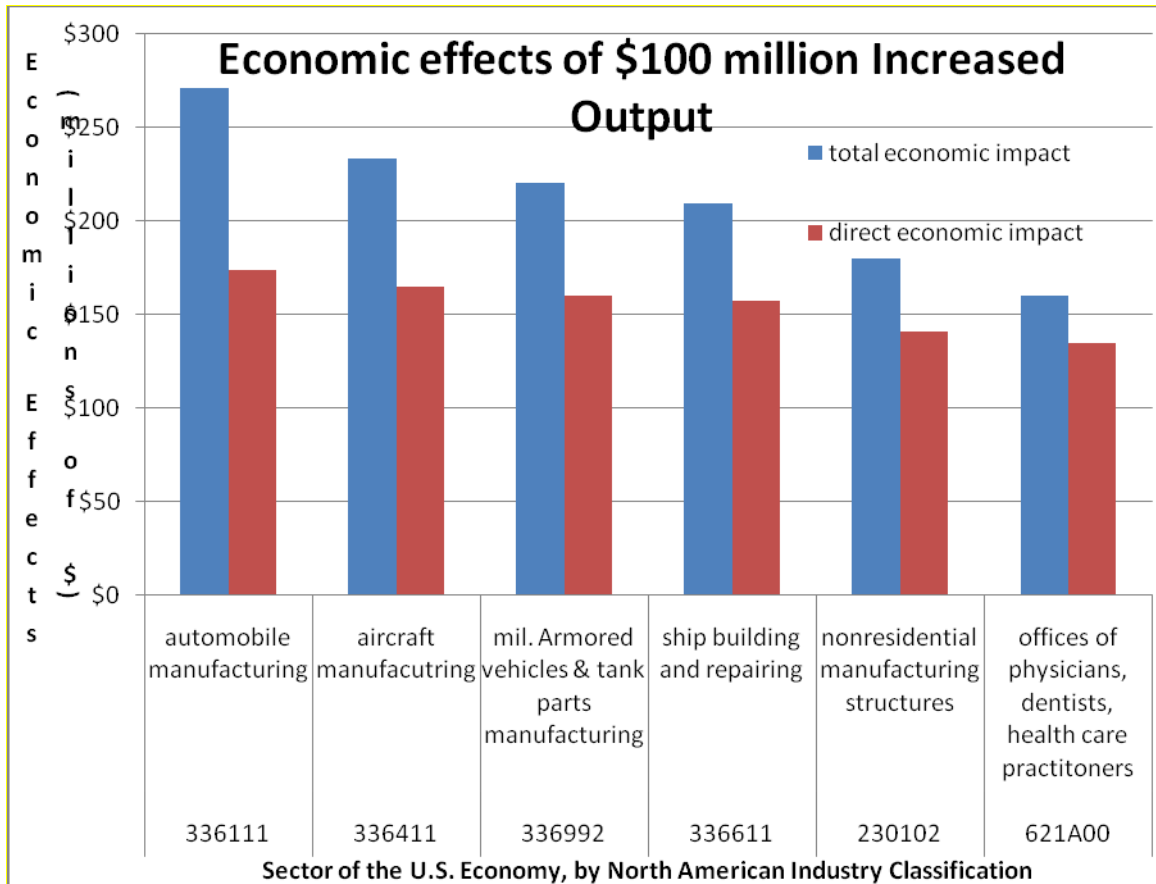


Figure 4. Economic Effects of \$100 Million Increased Output

The shipbuilding sector ranks third of the six sectors considered, when ranked by total economic effects, with a multiplier of 2.09 (\$209 million/\$100 million). Regional economic multipliers for each of the “big six” shipyards may actually be much higher.

Induced Economic Effects

Since \$209 million of total economic activity occurs for every \$100 million increased output from shipbuilding, the output multiplier, when considering only the impacts within the sector (direct) and the supply chain (indirect) is 2.09. Economists refer to this as a “type I multiplier” (Tessler, 2009). For every \$1 of increased output from shipbuilding, about \$2.09 of direct and indirect activity occurs.

What this analysis leaves out is the induced effects from the \$209 million of activity throughout the economy. The induced impacts are “employment and activity supported by the consumer spending of those employed in the sector or in its supply chain. This helps to support jobs in [US] industries that supply these purchases and includes jobs in retail outlets, companies producing consumer goods and in a range of service industries” (Oxford Economics, 2009). The induced multiplier is the most difficult to calculate or estimate, and the least credible for any industry or sector. Multipliers which include induced effects, economists call “type II multipliers” (Tessler, 2009).

Since the Carnegie-Mellon software does not include induced effects in generating the economic activity results, a type II multiplier is not explicitly calculated using the Leontief

inversion process. However, the constant that relates the “Type I and Type II multipliers in an input-output model [has been] proven to be exactly the consumption multiplier for the household sector” (Katz, 1980). The basic consumption multiplier is based on the Marginal Propensity to Consume (MPC) and the tax rate (t), and may be calculated as:

$$\frac{1}{1 - MPC(1 - t)}$$

An average tax rate of 35% was used, including federal income tax, social security, medicare, and possible state taxes. Assuming a national average marginal propensity to save (savings rate) of 7%, then $MPC = 1 - MPS = 93\%$. The result shows:

$$\frac{1}{1 - 0.93(1 - 0.35)} = 2.528$$

The result of using this consumption multiplier estimate to produce type II multipliers is included in Table 4 below.

Table 4.

<u>sector #</u>	<u>Sector</u>	(type I multiplier	minus increased demand)	* consumption multiplier	result -- type II multiplier
336111	automobile manufacturing	2.71	1	2.5	4.28
336411	aircraft manufacturing	2.33	1	2.5	3.33
336992	mil. Armored vehicles & tank parts manufacturing	2.2	1	2.5	3.00
336611	shipbuilding and repairing	2.09	1	2.5	2.73
230102	nonresidential manufacturing structures	1.8	1	2.5	2.00
621A00	offices of physicians, dentists, health care practitioners	1.6	1	2.5	1.50

In reality, differences exist in consumption multipliers between various sectors, but a realistic range is between 2.0 and 3.0, based on tax rates varying by region and MPC varying by profession or trade.

Labor

Using the Carnegie Mellon EIO-LCA Model

Using the Carnegie Mellon EIO-LCA model, the numbers below represent the complete number of employees needed across the supply chain of purchases in order to produce the level of output of \$100 million. The US economy-wide benchmark data used for this section is the 1997 benchmark data, since the 2002 model did not include the labor output functionality. *Here the shipbuilding and repair sector ranks 1st of the six considered, with 1670 additional employees needed throughout the supply chain in order to increase shipbuilding output by \$100 million.* Again, the model used is a linear model, so an increased output of \$1 billion would support 16,700 employees. The next most labor-intensive sector of the six considered is “military armored vehicles manufacturing,” which would utilize 1530 additional employees. ***Of the six specific sectors considered here for possible investments of federal government dollars (to increase that sector’s output),***



“shipbuilding and repair” will create or support the highest number of jobs. The results are shown in Figure 5 below.

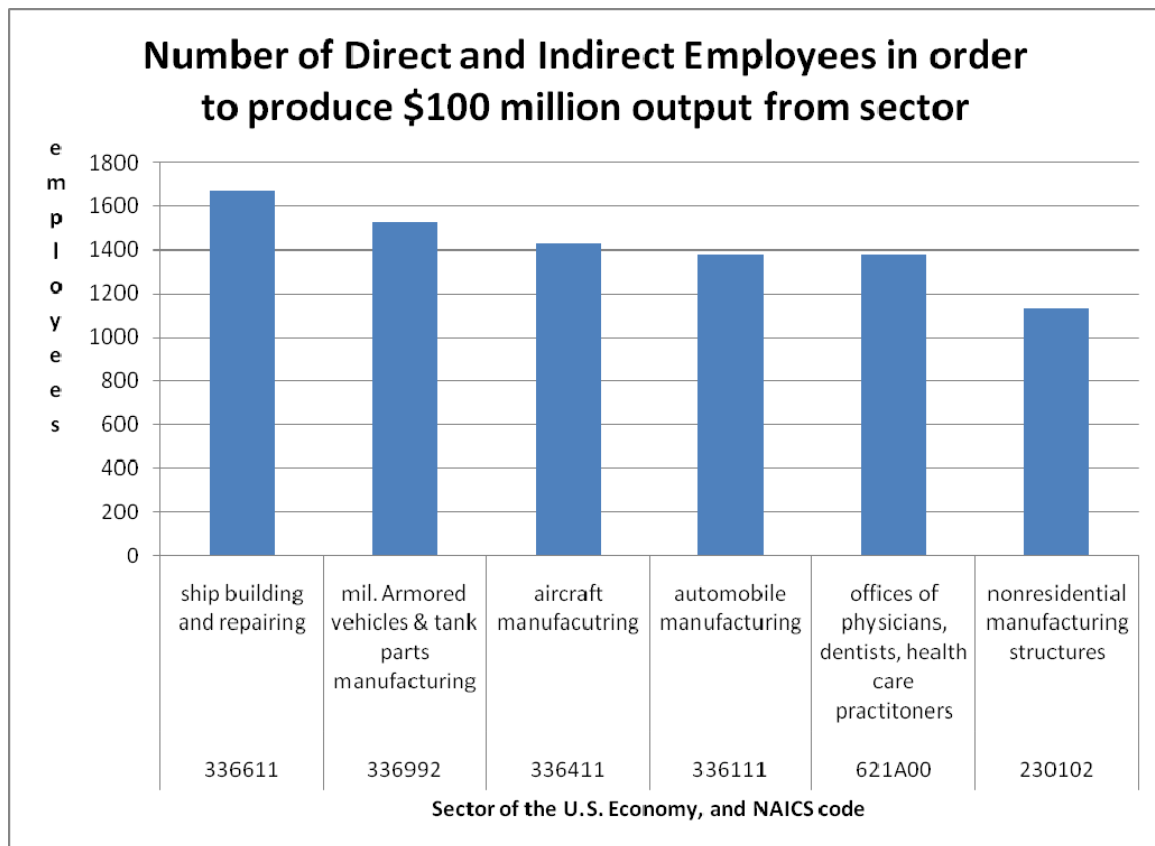


Figure 5. Number of Direct and Indirect Employees in Order to Produce \$100 Million Output from Sector

Of the 1,670 jobs created or supported by the shipbuilding and repair sector in order to create an additional \$100 million of output, the EIO-LCA model suggests that 918 of those jobs would be within the shipbuilding sector itself, while the remaining 752 would be throughout the supply chain (part of the indirect benefit).

Several assumptions and limitations are associated with the use of the EIO-LCA model to estimate increased employment based on a larger output demand. First, the data is old (1997 benchmark⁸). However, the industries selected are mature industries. Use of the model for information technology or telecommunications estimates would be much less reliable, as these sectors have experienced more widespread growth than shipbuilding, auto/aircraft manufacturing, or health care services. Secondly, the Bureau of Economic Analysis compiles benchmark data through surveys and forms submitted by US corporations to the federal government. Uncertainty in sampling, response rate, and errors in form completion are just a few of the potential sources of discrepancy between the data input and reality.

⁸ The Carnegie Mellon model does not support labor analysis for the 2002 benchmark data. The next most recent year for which benchmark input/output data is available from BEA is 1997.

In addition, the EIO-LCA is a producer price model—“the price a producer receives for goods and services (plus taxes, minus subsidies), or the cost of buying all the materials, running facilities, paying workers, etc.” The alternative pricing method, “purchaser price,” would include the producer price plus the transportation costs of shipping product to the point of sale, and the wholesale and retail trade margins (the profit these industries take for marketing and selling the product). For many goods, the producer prices can be far less than what a final consumer would pay (e.g., the producer price for leather goods in US is approximately 35% of the final purchaser price)” (Carnegie Mellon, 2008).

NAVSEA 05C Labor Data Trends

The “05C” office of Naval Sea Systems Command (NAVSEA) provided data for the employment of workers at private shipyards throughout the United States, over the last several decades. The data are presented in Figure 6, which shows a declining trend in the labor force levels at shipyards. As of the beginning of 2009, the nation’s “big 6” shipyards directly employed a total of over 56,000 workers. Twenty years ago, in January of 1990, the same yards employed over 80,000 people. In 2000, shipyard labor force levels were even lower than today, reaching about 67% of the 1990 levels. Today, 70% of the number of employees in 1990 are employed throughout the six private shipyards.

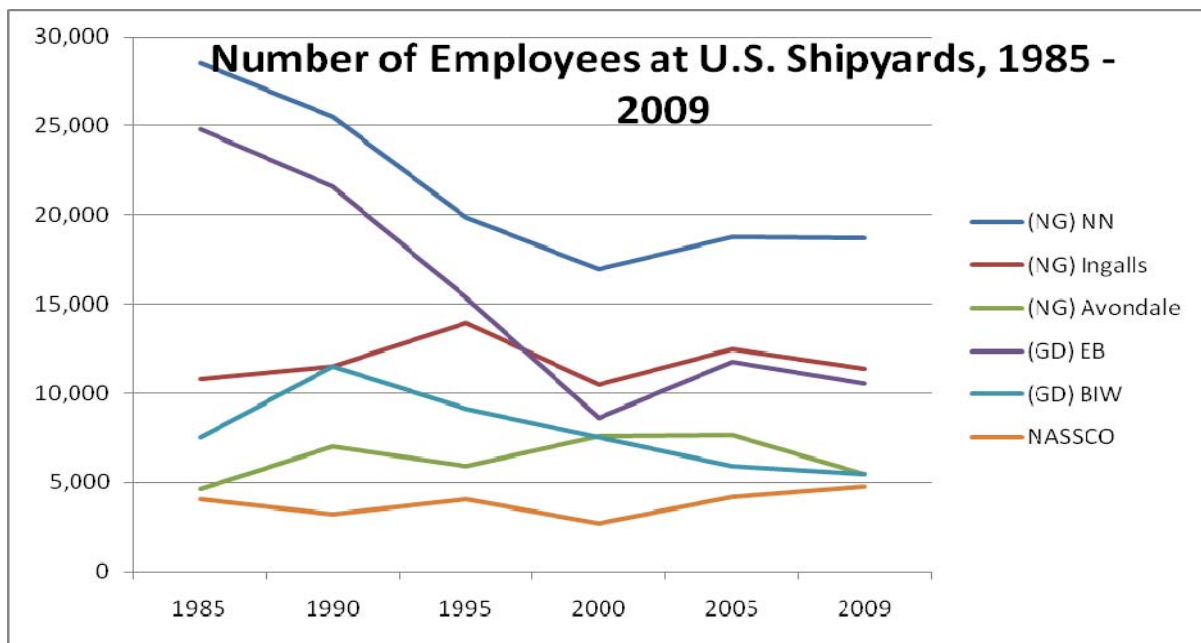


Figure 6. Number of Employees at US Shipyards, 1985-2009
(Source: NAVSEA 05C data file, last updated for all shipyards in 2009)

Figure 7 was produced by NAVSEA 05C's Portfolio Assessment Team, and shows the contribution to state GDP per shipyard worker, compared to the average worker in the state. The results show that in Maine, where Bath Iron Works employees over 5400 workers, on average they contribute more than nine times the income of an average Maine worker.⁹

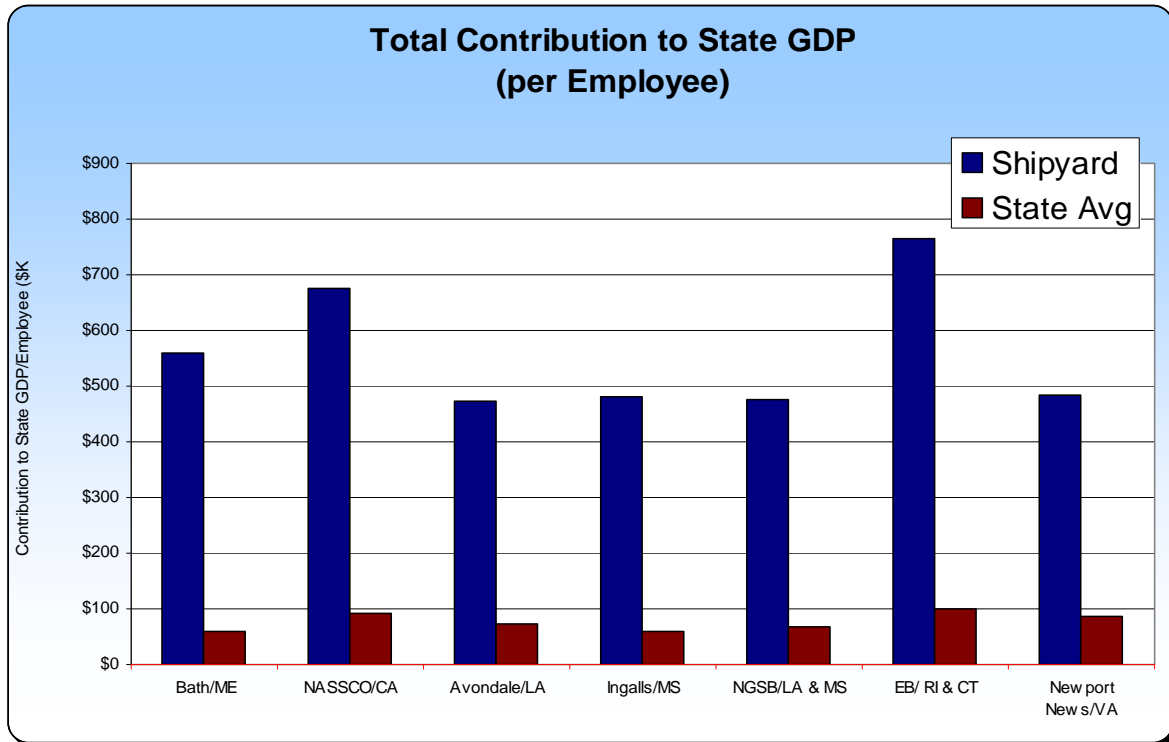


Figure 7. Total Contribution to State GDP

The results of the NAVSEA team's study suggest that shipyard workers contribute between 6 to 9 times more, on average per employee, than the average state worker.

Capital Intensity

Excess Capacity

The sections on Capital Intensity and Excess Capacity are still under development, and will be published upon completion of thesis research in May 2010. Thesis will be available online at <http://www.nps.edu/Library/>.

⁹ Source of average state wage data is the respective state's Bureau of Labor Statistics

References

- American Shipbuilding Association. (2009). *About ASA*. Washington, DC: Author.
- American Shipbuilding Association. (2009, December 3). Unilateral Navy disarmament. Washington, DC: Author.
- Atesoglu, H.S. (2005-6). Economic consequences of a rise in defense spending after September 11, 2001. *Journal of Post Keynesian Economics*, 181.
- Barnard, R.C. (2005). Shipbuilding: An uncertain future. *Sea Power*, 28.
- Booth, W., Colomb, G.G., & Williams, J.M. (2008). *The craft of research*. Chicago, IL: University of Chicago Press.
- Bureau of Economic Analysis. (2009, November 24). Gross domestic product release [Electronic Version]. Washington, DC: Author.
- Bureau of Economic Analysis. (2008, April). *Industry economic accounts*. Retrieved February 5, 2010, from http://www.bea.gov/industry/io_benchmark.htm
- Bureau of Labor Statistics. (2009). *Economy at a glance, United States*. Washington, DC: Author.
- Carnegie Mellon University Green Design Institute. (2008). *US 2002 industry benchmark model*. Retrieved February 3, 2010, from Economic Input-Output Life Cycle Assessment (EIO-LCA), <http://www.eiolca.net>
- Castelli, C. (2009, December 7). Navy raises 313-ship goal to 324, boosts focus on missile defense. *Inside the Navy*, p. 4.
- Department of the Navy. (2008). *Fiscal Year 2009 budget estimates*. Washington, DC: Author.
- Director, Warfare Integration (OPNAV N8F). (2010). *Report to Congress on annual long-range plan for construction of naval vessels for FY 2011*. Washington, DC: Office of the Chief of Naval Operations.
- Drake, R. L. (1976). A short-cut to estimates of regional input-output multipliers. *International Regional Science Review*, 1-17.
- Drucker, P. (1992). *The age of discontinuity: guidelines to our changing society*. Piscataway, NJ: Transaction Publishers.
- Ericson, M., He, E., & Schoenfield, A. (2009, June 24). Tracking the \$700 billion bailout. *The New York Times*, p. B2.
- Friedman, G. (2009). *The next 100 years*. New York: Doubleday.
- Gamick, D.H. (1970). Differential regional multiplier models. *Journal of Regional Science*, 35-47.
- Hancock, R. (2003). Underwater welding in nuclear power plants. *Welding Journal*, <http://www.aws.org/wj/sept03/feature1.html>.
- Hensel, D.N. (2009, November 17). Chief economist, United States Navy (Office of ASN, FM&C). [Interview with researcher].
- Homan, T. (2009, November 29). *Job losses, manufacturing probably eased: U.S. Economy Preview*. Retrieved December 11, 2009, from Bloomberg, <http://www.bloomberg.com/apps/news?pid=20601087&sid=aErLZXt7wkJs>
- Honolulu Advertiser. (2010, February 10). Pearl Harbor shipyard completes USS Hawaii repairs. *Honolulu Advertiser*.
- Katz, J.L. (1980). The relationship between type I and type II income multipliers in an input-output model. *International Regional Science Review*, 51-56.
- LECG, LLC. (2002). *Economic contribution of the U.S. commercial shipbuilding industry*. Washington, DC: Shipbuilders Council of America.
- LECG, LLC. (2002). *The economic contribution of the U.S. commercial shipbuilding industry*. Washington, DC: Shipbuilders Council of America.



- Leontief, W. (1966). *Input-output economics*. New York: Oxford University Press.
- Leontief, W. (1976). *Studies in the structure of the American economy*. White Plains: International Arts and Sciences Press.
- Lim, D. (2007). On the measurement of capital intensity. *Review of World Economics*, 760-766.
- McGuire, M. (2009). *Integration of the Naval Unmanned Combat Aerial System into the future Naval Air Wing*. Monterey, CA: Naval Postgraduate School.
- McIntire, K. (2009, September 16). Top admiral affirms commitment to 313-ship fleet. *Undersea Enterprise News Daily*, p. 2.
- Meyers, N.A. (2009). *The vitality of the vessels*. Monterey, CA: Naval Postgraduate School.
- Miller, R.E., & Blair, P.D. (1985). *Input-output analysis: Foundations and extensions*. Englewood Cliffs, NJ: Prentice-Hall.
- Mullen, A.M. (2007, 110th Congress 2nd session). Testimony before the Defense Appropriation Subcommittee.
- National Bureau of Statistics. (2009, October 10). *chinability.com*. Retrieved December 15, 2009, from GDP Growth in China, 1959-2009, <http://www.chinability.com/GDP.htm>
- National Defense Research Institute (RAND). (2006). *Why has the cost of Navy ships risen?* Washington, DC: RAND.
- Naval Historical Center. (2002, January 23). *U.S. Navy active ship force levels, 1917*. Retrieved December 15, 2009, from Naval Historical Center, <http://www.history.navy.mil/branches/org9-4.htm>
- NAVSEA Shipbuilding Support Office. (2009, December 1). *Fleet size*. Retrieved December 15, 2009, from Naval Vessel Registrar: <http://www.nvr.navy.mil/nvrships/FLEET.HTM>
- Northrop Grumman. (2009). *Shipbuilding fact sheet - Gulf Coast*. Pascagoula, MS: Northrop Grumman.
- O'Rourke, R. (2009). *Navy force structure and shipbuilding plans: Background and issues for Congress*. Washington, DC: Congressional Research Service.
- Oxford Economics. (2009). *The economic case for investing in the UK defence industry*. London: Oxford Economics.
- Oxford Economics. (2004). *The economic contribution of BAE Systems to the UK and implications for defence procurement strategy*. London: Oxford Economic Forecasting.
- Pace, J., Taylor, A., & Elliott, P. (2009, December 8). Obama urges major new stimulus, jobs spending. *The New York Times*, p. A1.
- Pershing, B. (2010, March 18). Congress sends Obama \$18 billion jobs bill to sign. *The Washington Post*.
- Praeger, M.A. (2009, December 3). VP, Congressional Affairs, American shipbuilding Association. [Interview with researcher].
- Recovery.gov. (2009). *The act*. Retrieved December 10, 2009, from Recovery.gov, http://www.recovery.gov/About/Pages/The_Act.aspx
- Redlick, R.J. (2009, October 1). Stimulus spending doesn't work. *The Wall Street Journal*, p. C5.
- Tessler, A.A. (2009, December 16). Economist, Oxford Economics, London Office (www.oef.com). [Interview with researcher].
- The World Bank. (2009, October 7). *World development indicators database*. Retrieved December 16, 2009, from Gross Domestic Product, 2008, <http://siteresources.worldbank.org/DATASTATISTICS/Resources/GDP.pdf>
- US Energy Information Administration. (2008, January). *U.S. energy information administration*. Retrieved December 14, 2009, from World Oil Transit Chokepoints, http://www.eia.doe.gov/cabs/World_Oil_Transit_Chokepoints/



- United States Congress. (2009, May 22). *Government printing office*. Retrieved December 15, 2009, from Weapons System Reform Act of 2009, <http://www.gpo.gov/fdsys/pkg/PLAW-111publ23/pdf/PLAW-111publ23.pdf>
- Walters, W. (2000). Geographic record: American naval shipbuilding. *Geographic Review*, 418-431.
- Work, R. (2009). *The U.S. Navy: Charting a course for tomorrow's fleet*. Washington, DC: Center for Strategic and Budgetary Assessments.
- Wright, J., & Fields, M. (2009). *Shipyard economic impact study*. Washington, DC: Naval Sea Systems Command.



Panel #15 – Applying Contemporary Economic Theory to Defense Acquisition Decision Making

Thursday, May 13, 2010

9:30 a.m. –
11:00 a.m.

Chair: Dr. James Simpson, Dean, College of Business Administration,
University of Alabama in Huntsville

***A Three Stage Multi-attribute Procurement Auction: A Proposal for
Department of Defense (DoD) Vendor Selection Decisions***

Francois Melese and Jay Simon, Naval Postgraduate School

***When More is Better: Design Principles for Prediction Markets in
Defense Acquisition Cost Forecasting***

Taroon Aggarwal, Ricardo Valerdi and Kevin Liu,
Massachusetts Institute of Technology; Jared Fortune, USC;
Matthew Potoski, Iowa State University; Marc
Giombetti, Technical University Munich, Germany; and John
Kim, GaN Corporation

***Innovations in Defense Acquisition: Implementing Information
Aggregation Markets***

Peter Coughlan, Bill Gates and Jeremy Arkes, Naval
Postgraduate School



A Three Stage Multi-attribute Procurement Auction: A Proposal for Department of Defense (DoD) Vendor Selection Decisions

Francois Melese—Dr. Melese joined NPS in 1982. Today he is Professor of Economics at the Defense Resources Management Institute (DRMI). In 2008, he helped edit the DoD's first Strategic Management Plan. Professor Melese has published over 50 articles and book chapters on a variety of topics and together with NPS colleagues was among the first to apply transaction cost economics to generate new insights into military cost estimating, and the make-or-buy decision. In 2009, NATO HQ asked Dr. Melese to organize a major NATO meeting to celebrate the 60th anniversary of the Alliance: "Building Integrity and Defense Institution Building."

Francois Melese
Defense Resources Management Institute (DRMI)
School of International Graduate Studies (SIGS)
Naval Postgraduate School
Code 64, NPS
699 Dyer Rd.
Monterey, CA 93943
fmelese@nps.edu
Office: 831-656-2009

Jay Simon—Dr. Simon is an Assistant Professor of Decision Science at the Defense Resources Management Institute (DRMI). His main research focus is multi-attribute preference modeling. His current and recent work includes a prostate cancer decision model, preference models for health decisions, preferences over geographic outcomes, altruistic utility modeling, and time discounting anomalies. He is a member of the Institute for Operations Research and the Management Sciences (INFORMS), and the Decision Analysis Society of INFORMS. Dr. Simon joined the DRMI faculty in August 2009.

Jay Simon
Defense Resources Management Institute (DRMI)
School of International Graduate Studies (SIGS)
Naval Postgraduate School
Code 64, NPS
699 Dyer Rd.
Monterey, CA 93943
jrsimon@nps.edu
Office: 831-656-2457
Home: 831-920-1454

Abstract

This paper explores vendor selection decisions where competing vendors offer similar products with multiple non-price attributes. Traditional "multi-attribute auctions" such as the USAF Air Tanker competition include prices alongside other attributes in vendor proposals. Buyers generally select winning bidders using a weighted average of price and non-price attributes. A different approach is recommended here: to embed vendor prices directly in the buyer's budget constraint. A first step is to conduct a simple multi-attribute auction with a fixed budget. The government buyer only evaluates vendor proposals that satisfy the budget constraint, choosing the proposal (non-price attribute bundle) which offers the greatest "value for money" (i.e., budget). The next step is to address budget uncertainty, expanding the model to incorporate a range of budgets. This leads to several interesting



results, including that the traditional practice, and classroom technique, of eliminating dominated alternatives can lead to sub-optimal decisions. Improving public procurement decisions requires forecasting a range of future budgets, and soliciting information from vendors that allows procurement alternatives to be defined as functions of the value offered by each vendor over a range of budgets, rather than as a single point in budget-value (cost-effectiveness) space. Under more realistic budget scenarios, different vendor selection decisions will occur that benefit both troops and taxpayers.

Background and Introduction

This paper offers a model to guide public procurement decisions in severely constrained budget environments. The global financial crisis swiftly evolved into a debt crisis that increasingly constrains government spending. The US is not immune. The federal debt, at roughly 80% of GDP and rising, is set to seriously constrain discretionary federal spending. Today, mandatory expenditures on Medicare, Medicaid, Social Security, Interest on the Debt, etc., make up over 60% of the federal budget, and continue to grow. By far the largest component of the corresponding, and shrinking, discretionary budget (more than half) is spent on national defense. The challenge then, is to develop an approach to accommodate the future budget uncertainties that will increasingly face federal agencies, and in particular, the US Department of Defense (DoD).

This paper explores vendor selection decisions where competing vendors offer similar products (computers, logistics packages, weapon systems, etc.) that incorporate multiple non-price attributes. In most multi-attribute auctions, vendor prices are included alongside other attributes as part of the vendor's proposal. Along with political considerations, a weighted average of price and non-price attributes is generally used by governments to help select the winning bidder. The approach recommended here is different. In the spirit of "Cost as an Independent Variable" (CAIV), the proposal is to incorporate vendor prices in the buyer's budget constraint.

A first step is to conduct a simple multi-attribute auction with a fixed budget (point estimate) for the program (product or service). In this case, the government buyer chooses from a set of vendors that each submit their best offers—non-price attribute bundles that fall within the specified budget. The next step is to address the political realities of budget uncertainty. This requires an expansion of the model that incorporates a range of possible budgets (e.g., optimistic, pessimistic, and most likely). This leads to several interesting results. Perhaps the most surprising is that the traditional practice, and classroom technique, of eliminating dominated alternatives can easily lead to sub-optimal decisions.

Finally, the model is generalized to capture budget uncertainty through the specification of a probability distribution (density function) over the range of possible budgets. The dual fundamental insights of this paper are that public procurement decisions can be improved if: i) interval estimates (ranges) of future budgets (and corresponding probability distributions) are forecasted, and ii) information from vendors is solicited that allows procurement alternatives to be defined as a function of value (non-price attributes) over a range of possible budgets, rather than as a single point in budget-value (cost-effectiveness) space.

Over the next five years, the US Department of Defense (DoD) plans to spend more than \$357 billion on the development and procurement of major defense acquisition programs (MDAPs). According to the Government Accountability Office (GAO, 2009), the DoD's goal is to "achieve a balanced mix of weapon systems that are *affordable*" [emphasis added]. With this goal in mind, a multi-attribute first price, sealed bid procurement auction is



proposed that extends traditional price-only auctions to one in which competition takes place exclusively over specific bundles of desired non-price attributes or characteristics.

In this model, prices/costs do not appear in the buyer's value function. Instead, in the spirit of "cost as an independent variable" (CAIV), prices/costs are incorporated in the buyer's affordability constraint. Larsen (2007) provides the following explanation of CAIV:

All acquisition programs/issues consist of three fundamental elements: cost, performance and schedule. Under CAIV, performance and schedule are considered a function of cost. Cost and affordability should be a driving force not an output after potential solutions are established.

Michael and Becker (1973) and others also discuss the importance of separating prices from multi-attribute measures of value. This is especially important for agencies that face budget uncertainty.

For example, given the growing pressure on defense budgets from expanding federal deficits, the military's "bow wave" of planned procurement contracts and operating cost commitments is not affordable, and has recently threatened several MDAPs. [e.g., the Army's Future Combat System (FCS) and the Air Force Joint Strike Fighter (JSF) program] An important lesson is that different optimal vendor selections likely would have taken place under more realistic budget scenarios, and that this would have benefited both troops and taxpayers.

This paper explores vendor "performance competition" under alternative budget scenarios. Given a target budget for the program (e.g., computers, vehicles, weapons, logistics packages, etc.), competing vendors generate multi-attribute (performance) offers based on their individual costs, technology, productivity, production processes, supply chains, etc. Each vendor (seller/bidder) is provided the same budget authority guidance from the procurement agency's projected funding forecast (a point estimate). Vendors respond by offering their best possible non-price attribute bundles.

In the model, a vendor's proposal (offer) depends on the buyer's budget, and the individual vendor's costs of supplying each attribute along with the production technology they have available to combine those attributes. Competition between vendors takes place exclusively over product performance (collections of non-price attributes) for a given budget. The government buyer's value function facilitates the evaluation vendor proposals (bundles of non-price attributes) and the selection of a winner.¹⁰

¹⁰ Consider a simple example. Suppose a dozen impact resistant notebook computers are required for a provincial reconstruction team (PRT) heading to Afghanistan. The team believes the most likely budget available for these computers will be \$20,000, and it reveals this to the competing vendors. For this particular budget, each vendor offers a different combination of components, attributes and quality characteristics based on its technology, its capital and labor productivity, and the fixed and variable costs embedded in its supply chains. Given the different sets of a dozen computers offered by the each vendor, the team then selects the vendor that offers the best value (mix of attributes) for their budget, or the one that maximizes their private value function. With the funding uncertainty inherent in defense programs, this paper suggests that vendors be asked to provide offers for more than one budget level (say an optimistic budget of \$25,000 and a pessimistic budget of only \$15,000). In general, the recommendation is that each competing vendor's proposal be re-defined as a function of the bundle of multiple (non-price) attributes they can offer over the range of possible budget levels (e.g., optimistic, pessimistic, and most likely).

A straightforward extension is to allow the buyer to offer a range of possible budgets. However, this involves a more complex solicitation, inviting proposals from different vendors for each possible budget. The significant benefit is that it provides a more robust view of a vendor's ability to provide performance. The "expansion path" generated for each vendor shows how that vendor's proposal changes as the budget increases or decreases.

This approach can be thought of as a strategic choice of auction mechanism for a buyer when possible overall budget authorities for the program can only be estimated/forecasted, and the products/services are highly differentiated and complex. It combines the competitive advantages of auctions with the flexibility of a decision based on multiple attributes of the product.

In our formulation, both seller and buyer suffer from imperfect and asymmetric information. The seller does not know the relative weights the buyer assigns to the attributes. Meanwhile, the buyer does not know the sellers' costs of producing a particular attribute, nor the technology (production functions) that combines those attributes into the desired products. Parkes and Kalagnanam (2005) explain this asymmetry in the case of a seller's private information: "[S]eller costs can be expected to depend on [the] local manufacturing base and sellers can be expected to be well informed about the cost of (upstream) raw materials."

Loerch, Koury, and Maxwell (1999) discuss a "Value Added Analysis" approach that is similar to ours in employing multi-attribute preferences in weapon systems acquisition decisions. Our approach differs from theirs in that we incorporate the vendors' decision-making explicitly into the model, and capture the issue of asymmetric information.

Blondal (2005) discusses a two-stage¹¹ bidding process similar to ours, in which the procuring agency issues a general request, and then later issues a more detailed request based on the responses received. The US Federal Acquisition Regulations (2005) provide guidance in subpart 14.5 on another two-step process for government agencies:

Step one consists of the request for, submission, evaluation, and (if necessary) discussion of a technical proposal. No pricing is involved. Step two involves the submission of sealed price bids by those who submitted acceptable technical proposals in step one. Invitations for bids shall be issued only to those offerors submitting acceptable technical proposals in step one. An objective is to permit the development of a sufficiently descriptive and not unduly restrictive statement of the Government's requirements especially useful for complex items.

The approach proposed in this paper is similar, but differs from this two-step bidding process in that the competition is over non-price attributes, and the price is captured in the budget authority (or affordability) constraint.

Much of the multi-attribute auction literature (Che, 1993; Beil & Wein, 2003; and Parkes & Kalagnanam, 2005; etc.) either implicitly or explicitly includes price alongside non-

¹¹Blondal defines a "stage" differently than we do in this paper. We use the term to refer to a decision or set of decisions that depends only on exogenously given parameters and previous decisions. For example, Blondal considers a government agency's solicitation and the vendor offers in response to be a single stage, whereas we treat these as two distinct stages. Using our interpretation, Blondal's model is in fact a five-stage process.

price attributes in the buyer's (auctioneer's) value/utility function.¹² While this approach is appropriate in some contexts, it can generate complications in evaluating alternative defense investments. Unlike private sector decision makers that maximize profits (Revenues minus Costs), the government cannot simply subtract prices (or costs) from non-price attributes and generate an equivalent “profitability” metric to evaluate alternative investment options.

Interestingly, a necessary (but not sufficient) condition for a firm to maximize profits is for it to maximize the value of its output for the costs it incurs (or in the dual: to minimize the costs of producing its output). Interpreting the firm's costs as its budget, this corresponds to maximizing output for a given budget, and is analogous to our proposal for a government activity to maximize its value function (of non-price attributes) for a given affordability (budget) constraint. In fact, the proposal is to promote an approach where government decision makers maximize their value function (over a set of non-price attributes) for a range of budgets they think they might be allocated, but that might not be revealed with certainty within the window of that decision process.

However, historically (and even today) budgets for specific defense products and services were generated by the low cost (or best value) vendor among those responding with price quotes, for the requirements set by the military branches to support warfighters. Given current fiscal deficit projections that promise increasingly tighter budget constraints, it is not likely this approach will survive much longer. Traditional “bottom-up” budgets historically generated by the military services in cooperation with vendors to purchase force requirements, are likely to be subjected to greater top-down budget guidance.¹³

Two widely cited pioneers in defense economics, Hitch and McKean (1967), advocate an approach similar to that proposed in this paper: To determine the “maximum effectiveness for a given budget” and to examine how each alternative fares for several different budget levels.

The Model

The model consists of three stages, illustrated in Figure 1.

¹²Note that in defense procurement, value functions generate “measures of effectiveness” (MOEs). The term “MOE” is used in a few different ways. It may describe a single-attribute value function, or a multi-attribute value function which might incorporate the whole objectives hierarchy, or only a portion of it. For a detailed discussion of MOEs, see Sproles (2000).

¹³ Some evidence of tighter fiscal guidance appears in the new emphasis on fiscally informed “Planning” that involves more up-front military investment trade-offs, and correspondingly stricter fiscal guidance to the Military Services, in DoD's over-arching “Planning, Programming, Budgeting and Execution System” (PPBES). (See the DoD Comptroller's website.)

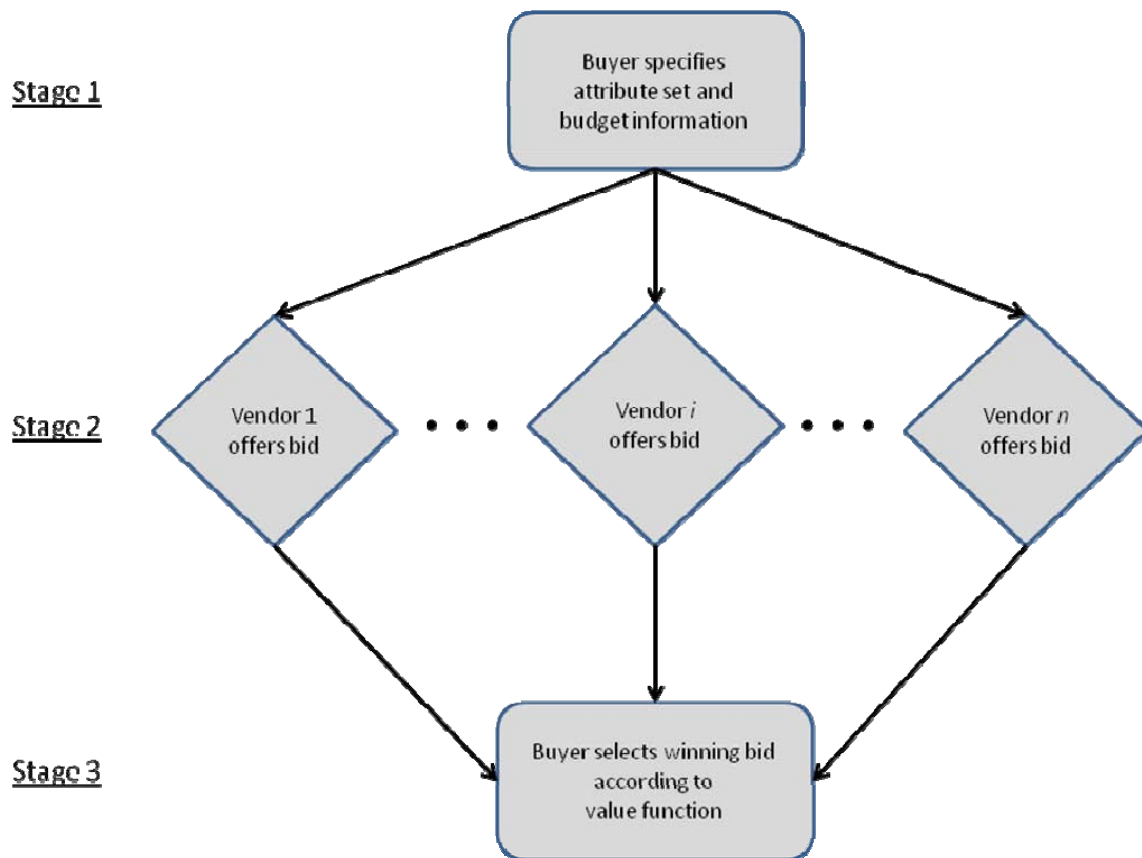


Figure 1. The Three-stage Procurement Model

In the first Stage, the procurement agency (buyer) solicits offers from vendors (sellers), specifying a set of attributes A and a budget level B . There are n vendors, each of whom responds in the second stage with a bid. In this case, a “bid” is simply a set of non-price attribute levels a vendor offers to produce for the budget, B . We express vendor i ’s bid as $A_i = [a_{i1}, \dots, a_{im}]$ for $i = 1, \dots, n$, where a_{ij} is the level of attribute j offered by vendor i . In the third stage, the buyer’s decision is to select a vendor, $i \in [1, n]$, and thus a set of attribute levels $A_i = [a_{i1}, \dots, a_{im}]$, that maximizes the “measure of effectiveness” (MOE), which we express as the value function $V(A_i)$. We assume $V(A_i)$ is an additive multi-attribute value function, though as we will observe later, our conclusions do not require $V(A_i)$ to be additive. The use of additive multi-attribute value functions requires that preferential independence be satisfied (Dyer & Sarin, 1979; Kirkwood & Sarin, 1980)¹⁴. We assume for simplicity that the single-attribute value functions are linear, and that attributes are

¹⁴ See Keeney and Raiffa (1976) and Kirkwood (1997) for further discussion of additive multiattribute value functions. For discussion of the elicitation and use of such preferences in defense applications, see Parnell (2007).

measured on the same scale¹⁵. Thus, we will refer simply to a_{ij} rather than using the notation $v(a_{ij})$. The buyer's objective is:

$$\max_i V(A_i) = \sum_{j=1}^m w_j a_{ij}, \quad (1)$$

where w_j is the weight the buyer places on attribute j . We assume the buyer has an understanding of the range of attribute levels when determining the weights, and that these weights are private information to the buyer. Asker and Cantillon (2007) refer to this as a "secret scoring rule." The final stage of the model is the application of (1) to the set of bids, and the selection of the vendor yielding the highest value.

Given A and B , each vendor chooses an attribute bundle which meets the budget constraint revealed by the buyer. A vendor has private information regarding production capabilities and costs, but must somehow form beliefs about the likelihood of a bid being accepted. We facilitate formulation of these beliefs by having each vendor generate a "best guess" at the weights of the buyer's (additive) value function, which we can express as $W_i = (w_{i1}, \dots, w_{im})$. We refer to this hypothetical value function as $Q(A_i)$. A higher $Q(A_i)$ indicates a greater probability with which the vendor believes the bid will be accepted. Only the ordinal rankings imposed by this value function are relevant, since vendors in our model will simply choose the attribute bundle (s)he believes has the highest probability of being chosen.

We can express the problem faced by vendor i as:

$$\begin{aligned} \max_{a_{ij}} Q(A_i) &= \sum_{j=1}^m w_{ij} a_{ij}, \quad j = 1, \dots, m \\ \text{s.t. } TC_i &= \sum_{j=1}^m c_{ij}(a_{ij}) \leq B, \end{aligned} \quad (2)$$

where the total cost TC_i is an additive function of the costs of firm i to produce each attribute level. The total costs for a particular vendor of generating its (non-price attribute bundle) offer cannot exceed B .

The individual attribute cost functions are given by $c_{ij}(a_{ij})$, and each one is increasing in a_{ij} . Because the objective function of (2) is linear, a unique solution exists, provided that $c_{ij}(a_{ij})$ is strictly convex for $j = 1, \dots, m$. This is reasonable, since it simply corresponds to decreasing returns to scale from investment in improving an attribute level.

For purposes of illustration and ease of exposition, the remainder of the study focuses on only two vendors, and two (non-price) attributes measured on the same scale. Assuming each vendor has a different technology (production process) to combine the two

¹⁵The structure of the single-attribute value functions is not germane to the purpose of this paper. The same results can be obtained without these assumptions, but we believe that some of the subsequent examples will be more illustrative when shown in attribute space rather than value space. Some clarity might be provided by thinking of the attribute levels as performance ratings.

attributes and faces different attribute cost functions, the Lagrangian function for the vendor's problem is given by:

$$L_i = w_{i1}a_{i1} + w_{i2}a_{i2} - \lambda_i (B - c_{i1}(a_{i1}) - c_{i2}(a_{i2})), \text{ for } i = 1, 2. \quad (3)$$

Since $\frac{\partial Q(A_i)}{\partial a_{ij}} > 0$ for both attributes, we can assume each vendor will use the maximum available budget B to produce the attribute bundle proposal. The first order necessary conditions for an optimum are given by:

$$\frac{\partial L_i}{\partial a_{i1}} = w_{i1} + \lambda_i c'_{i1}(a_{i1}) = 0 \quad (4a)$$

$$\frac{\partial L_i}{\partial a_{i2}} = w_{i2} + \lambda_i c'_{i2}(a_{i2}) = 0 \quad (4b)$$

$$\frac{\partial L_i}{\partial \lambda_i} = B - c_{i1}(a_{i1}) - c_{i2}(a_{i2}) = 0, \quad (4c)$$

where 4(c) simply asserts that the entire budget is being used. Solving (4a) and (4b) yields:

$$\frac{w_{i1}}{c'_{i1}(a_{i1})} = \frac{w_{i2}}{c'_{i2}(a_{i2})}, \quad (5)$$

meaning each vendor should choose a bid that uses the entire budget, and for which the two attributes have equal ratios between the (subjective) belief of the weight placed on the attribute by the buyer, and the vendor's private marginal costs¹⁶. The solicitation for bids results in two vendor offers (a_{11}, a_{12}) and (a_{21}, a_{22}) for the buyer to evaluate. The buyer simply selects the vendor whose bid maximizes V .

In general, w_{ij} and c_{ij} , are likely to differ between vendors. Multi-attribute auctions allow vendors to differentiate themselves in the auction process and to bid on their competitive advantages (Wise and Morrison 2000).

Multiple Budgets and Expansion Paths

With the preliminary model in place, the next step is for the buyer to more fully explore differences between vendors. Rather than the buyer specifying a budget B , the buyer now specifies a set of (increasing) possible budgets: B_1, \dots, B_k . Each vendor will go through the process described in section 2, k times, and produce a bid satisfying (5) for each of the k possible budgets. This set of bids from a vendor constitutes an "expansion path." It reveals to the buyer precisely how a vendor's bid will improve as the budget increases.

For purposes of illustration, it is helpful to consider a particular functional form:

¹⁶Note that (5) has a unique solution for each vendor when the entire budget is being used. Since both cost functions are increasing and strictly convex, as we move along the budget constraint curve, one marginal cost is decreasing and the other is increasing.

$$c_{ij}(a_{ij}) = \alpha_{ij} e^{\beta_{ij} a_{ij}}, \quad \alpha_{ij}, \beta_{ij} > 0 \text{ for } i=1,2, \quad j=1,2. \quad (6)$$

Note that the functions described by (6) are strictly increasing and convex. The exponent β_{ij} defines the convexity of each cost function. Although the insights and conclusions that follow do not depend on this particular functional form, it simplifies the analysis to use these cost functions throughout the remainder of the paper.

There are three reasons the expansion paths observed by the buyer might differ between vendors: i) the parameters of their cost functions could differ (α_{ij}, β_{ij}), ii) their beliefs about the buyer's value function could differ (w_{ij}), or iii) both the cost functions and the beliefs could differ.

First, consider the case in which both vendors believe the buyer places equal weight on the two attributes, but the vendors differ in their capabilities of producing those attributes. Specifically, suppose:

$$\alpha_{11} = \alpha_{12} = 2.0, \beta_{11} = \beta_{12} = 0.6, \alpha_{21} = \alpha_{22} = 1.0, \beta_{21} = \beta_{22} = 1.0, w_{11} = 0.7, w_{21} = 0.7. \quad (7)$$

Note that (7) reflects symmetry between the two attributes in the sense that neither vendor “specializes” in producing a particular attribute. For simplicity, we use cost functions with these properties throughout this section, and also to emphasize that asymmetry between attributes is not required to realize the benefits of this expansion path approach. Applying (7) results in the expansion paths shown in Figure 2.

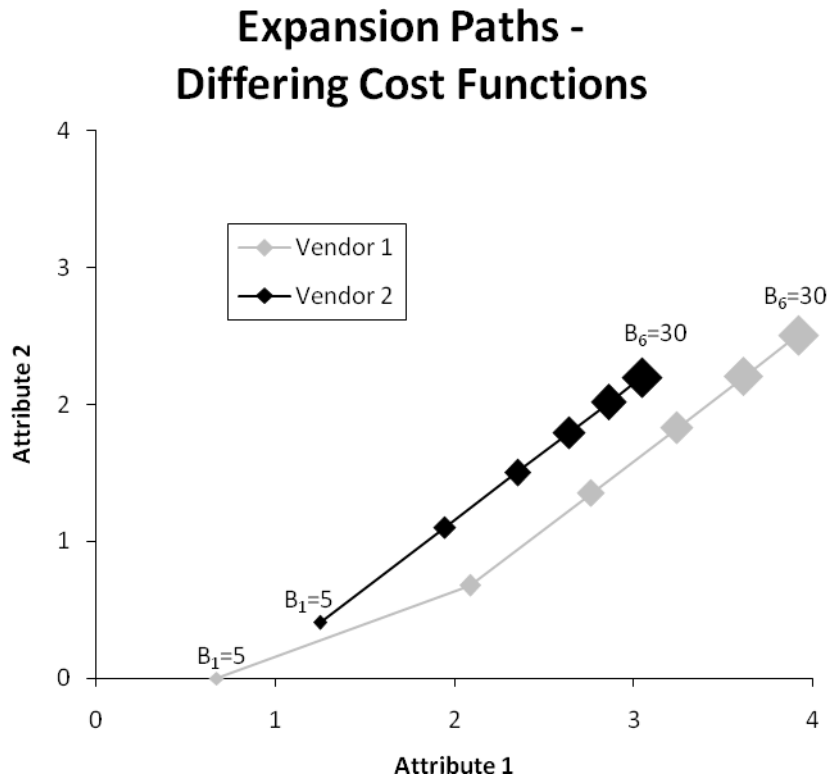


Figure 2. Expansion Paths for Two Vendors with Differing Cost Functions as the Budget Increases from 5 to 30 (the Markers of Increasing Size Show Each Vendor's Proposal as the Budget Increases in Increments of 5)

Alternatively, vendors could face the same cost functions, but express different beliefs about the buyer's attribute weights. Specifically, suppose the vendors have the following parameter values:

$$\alpha_{11} = \alpha_{12} = \alpha_{21} = \alpha_{22} = 2.0, \beta_{11} = \beta_{12} = \beta_{21} = \beta_{22} = 0.6, w_{11} = 0.5, w_{21} = 0.7. \quad (8)$$

That is, vendor 2 believes the buyer will place a slightly greater weight on attribute 1, while vendor 1 believes the weights on attribute 1 and attribute 2 will be equal. This results in the expansion paths shown in Figure 3.

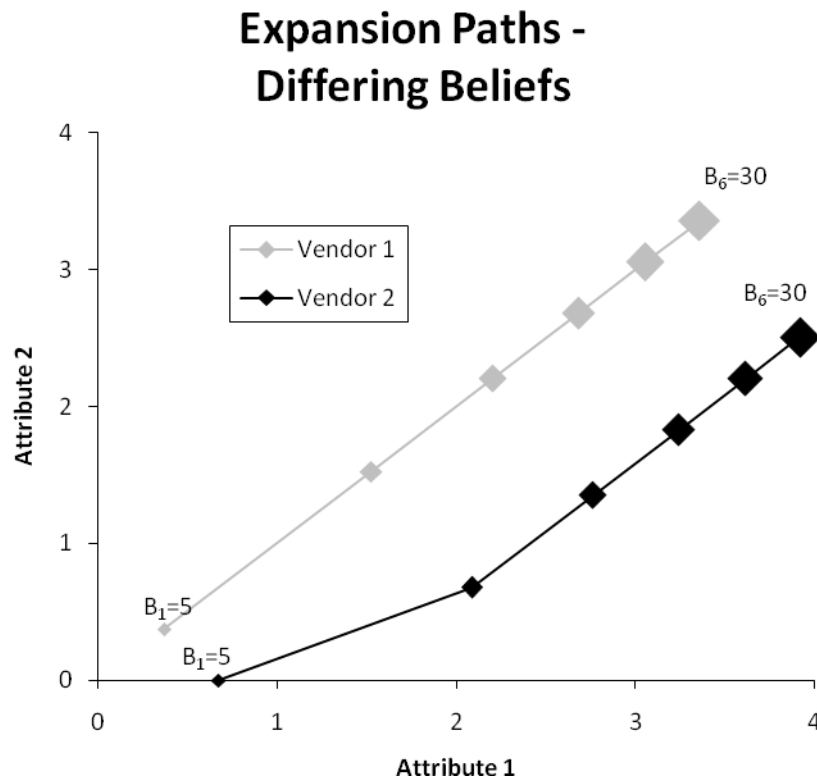


Figure 3. Expansion Paths for Two Vendors with Differing Beliefs on Attribute Weights as the Budget Increases from 5 to 30 (the Markers of Increasing Size Show Each Vendor's Proposal as the Budget Increases in Increments of 5)

While (7) and (8) are interesting special cases, it is also possible the two vendors will differ in both their costs and their beliefs. Consider two vendors with parameter values:

$$\alpha_{11} = \alpha_{12} = 2.0, \beta_{11} = \beta_{12} = 0.6, \alpha_{21} = \alpha_{22} = 1.0, \beta_{21} = \beta_{22} = 1.0, w_{11} = 0.5, w_{21} = 0.7. \quad (9)$$

In this case, we observe an interesting dynamic, as shown in Figure 4.

Expansion Paths - Differing Beliefs and Cost Functions

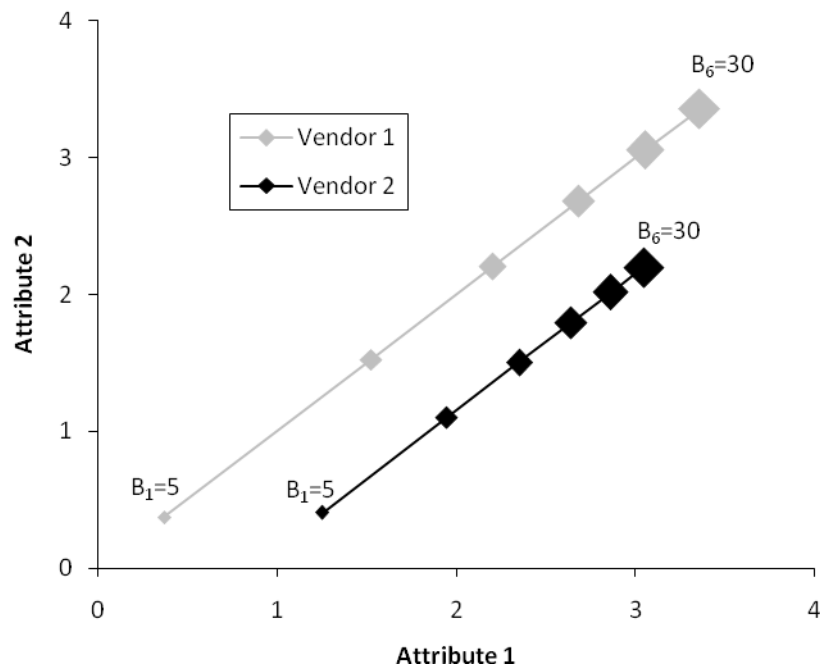


Figure 4. Expansion Paths for Two Vendors with Differing Costs and Beliefs on Attribute Weights (the Markers of Increasing Size Show Each Vendor's Proposal as the Budget Increases in Increments of 5)

For a high budget (e.g., 20), vendor 1 dominates vendor 2. Regardless of the buyer's preferences (provided value is monotonically increasing in each attribute), (s)he will select vendor 1. In a static comparison that assumes a fixed budget of 20, vendor 2 would be eliminated from further consideration. However, if the buyer must proceed knowing that a budget cut is possible, a dominated alternative may in fact become the preferred one. The reverse phenomenon can also occur. Notice in Figure 4 that vendor 2 dominates vendor 1 at a budget level of 5. A static comparison assuming a fixed budget of 5 would eliminate vendor 1 from further consideration.

To more clearly illustrate this phenomenon, we first assign attribute weights to the buyer's value function. Let the buyer assign a weight of 0.7 to attribute 1, and 0.3 to attribute 2. Instead of operating as before in attribute space, we now plot the two vendors' bids illustrated in Figure 4 as curves in "budget-value" space:

Value by Budget Level

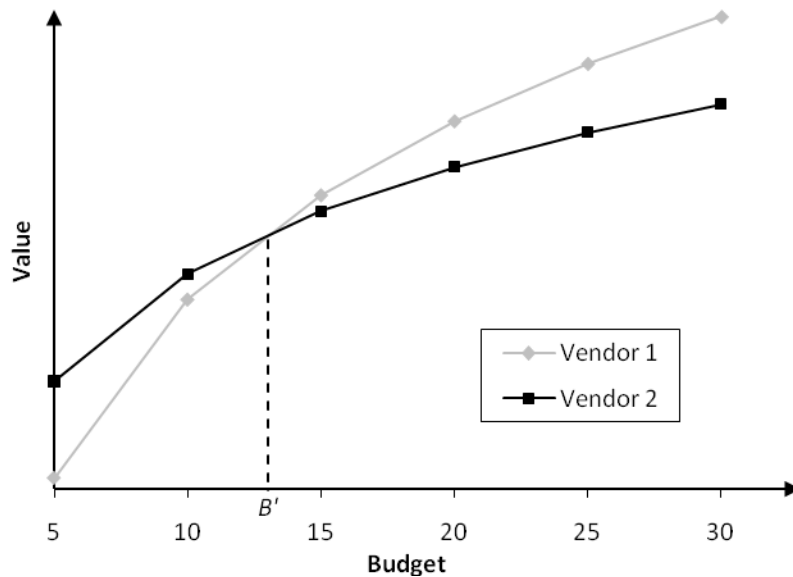


Figure 5. Value Provided by Each Vendor's Bid for Various Budget Levels

It is clear from Figure 5 that vendor 2 dominates the competition for any positive budget below the switch-point, $B < B'$, while vendor 1 dominates for any budget above the switch-point, $B > B'$. This suggests rethinking the typical definition of dominance in the literature, a concept routinely used in classroom illustrations and real-world applications to eliminate vendors that refers to points (not functions) in cost-effectiveness space.

In fact, viewing alternative vendors as functions in budget-value space reveals that the traditional definition of dominance can be misleading. For example, consider offers from vendor 1 and vendor 2 based on a very optimistic budget above B' . The traditional technique that focuses on points and not functions would likely eliminate vendor 2. Yet it is clear from Figure 5 that eliminating vendor 2 prematurely could lead to a less desirable outcome if the budget turned out to be wildly optimistic and the real budget was actually somewhere in the range of $0 < B < B'$.

This phenomenon can occur whenever any two vendors' expansion paths are shaped differently, and there is nothing unique about the particular functions chosen in our example; they were only selected for ease of exposition. Moreover, the same result can easily occur in the case of non-additive forms of the buyer's value function, and any non-linear interactions between attributes is likely to further magnify the effect¹⁷. Our conclusion is that this is a basic result that can arise in a wide variety of defense acquisition decision environments. Addressing this issue explicitly with this simple new approach could greatly benefit both our troops and taxpayers. Given the growing US federal deficit, future budget

¹⁷For example, consider a multiplicative value function for the buyer, and suppose that one vendor has to incur a large cost to provide anything above the minimum level for one particular attribute. This vendor will offer bids of little to no value for low budgets, but depending on cost functions and beliefs, may offer very attractive bids for higher budgets.

challenges may make it imperative for our government, and in particular the Department of Defense, to adopt an approach similar to the one proposed in this study.

Conclusion

We have described a simple three-stage multi-attribute procurement process for defense acquisitions. It allows the buyer to incorporate preferences over multiple attributes, and it allows each vendor to offer their best possible bid based on the budget provided, and the vendor's private cost structure and beliefs. Unlike most current methods, our model applies the spirit of CAIV; we do not include costs or any related price attributes in the buyer's value function. Instead, cost enters the model as part of a budget constraint.

The basic model is easily extended to allow vendors to submit bids for multiple potential budget levels. This leads to the generation of expansion paths for each vendor, which illustrates to the buyer precisely how a particular vendor's bid will improve under more optimistic budgets or slip under progressive budget cuts.

Interestingly, it can easily turn out that a vendor whose bid is "dominated" at an optimistic budget level is actually the most desirable choice at a more realistic (pessimistic) budget level. As a consequence it is vital for procurement agencies to reset their vendor evaluations, and to begin viewing each alternative (vendor) as a curve in budget-value space, rather than as a single cost-effectiveness point. Given future budget realities, this expanded view will prevent a vendor from being prematurely eliminated from consideration when budgets are likely to change over time.

The dual fundamental insights of this study are that public procurement decisions can be improved if: i) interval estimates (ranges) of future budgets (and corresponding probability distributions) are forecasted, and ii) information from vendors is solicited that allows procurement alternatives to be defined as a function of value (non-price attributes) over a range of possible budgets, rather than as a single point in budget-value (cost-effectiveness) space. The key implication is that different vendor selection decisions are likely to occur under more realistic budget scenarios, and that this can benefit both troops and taxpayers.

References

- Asker, J., & Cantillon, E. (2007). Procurement when price and quality matter. *CEPR Discussion Papers 6082*. London: Centre for Economic Policy Research.
- Beil, D., & Wein, L. (2003). An inverse-optimization-based auction mechanism to support a multiattribute RFQ process. *Management Science*, 49(11).
- Che, Y. (1993). Design competition through multidimensional auctions. *RAND Journal of Economics*, 24(4), 668-680.
- Dyer, J.S., & Sarin, R.K. (1979). Measurable multiattribute value functions. *Operations Research*, 27(4), 810-822.
- GAO. (2009, March 18). Testimony of M. Sullivan, Director of Acquisition and Sourcing Management before the Committee on the Budget, House of Representatives. Washington, DC.
- GSA, DoD, NASA. (2005, March). *Federal acquisition regulation (FAR)*. Washington, DC: Author.



- Hitch, C., & McKean, R. (1967). *The economics of defense in the nuclear age*. Cambridge, MA: Harvard University Press.
- Keeney, R.L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value trade-offs*. New York: John Wiley and Sons, and Cambridge University Press.
- Kirkwood, C. (1997). *Strategic decision making*. New York: Duxbury Press.
- Kirkwood, C.W., & Sarin, R.K. (1980). Preference conditions for multiattribute value functions. *Operations Research*, 28(1), 225-232.
- Loerch, A.G., Koury, R.R., & Maxwell, D.T. (1999). Value added analysis for army equipment modernization. *Naval Research Logistics*, 46(3), 233-253.
- Michael, R. T., & Becker, G.S. (1973). On the new theory of consumer behavior. *The Swedish Journal of Economics*, 75(4), 378-396.
- Parkes, D., & Kalagnanam, J. (2005). Models for iterative multiattribute procurement auctions. *Management Science*, 51(3).
- Parnell, G.S. (2007). Value-focused thinking using multiple objective decision analysis. In A. Loerch & L. Rainey (Eds.), *Methods for conducting military operational analysis: Best practices in use throughout the Department of Defense*. Alexandria, VA: Military Operations Research Society.
- Sproles, N. (2000). Coming to grips with measures of effectiveness. *Systems Engineering*, 3(1), 50-58.
- Wise, R., & Morrison, D. (2000). Beyond the exchange—the future of B2B. *Harvard Business Review*, 78(6), 86-96.

When More is Better: Design Principles for Prediction Markets in Defense Acquisition Cost Forecasting

Taroon Aggarwal—Taroon Aggarwal is a graduate student at MIT, doing his Masters in Systems Design and Management. He is also working as a research assistant at MIT Lean Advancement Initiative (LAI). His research interests lie in the areas of cost estimation, software engineering and application of collective intelligence tools to large organizations. Prior to his graduate studies, he worked for around five years as an IT consultant in the financial services domain. Taroon also holds a Bachelor's degree in engineering from India.

Taroon Aggarwal
Research Assistant
Lean Advancement Initiative
Massachusetts Institute of Technology
tar_agg@mit.edu

Ricardo Valerdi—Ricardo Valerdi, PhD, is a Research Associate at MIT. He received his doctoral degree in systems engineering from USC in 2005, where he developed the COSYSMO model for systems engineering cost estimation. He is also a Senior Member of the Technical Staff at the Aerospace Corporation. Prior to his doctoral studies, Dr. Valerdi was a systems engineer at Motorola Radio Network Solutions Group. He served on the INCOSE Board of Directors as Treasurer and is co-Editor-in-Chief of the *Journal of Enterprise Transformation*.

Dr. Ricardo Valerdi
Research Associate
Lean Advancement Initiative
Center for Technology, Policy, and Industrial Development
Massachusetts Institute of Technology
rvalerdi@mit.edu

Matthew Potoski—Matthew Potoski is an Associate Professor in the Department of Political Science at Iowa State University, where he teaches courses on public management and policy. He has received Iowa State University LAS awards for Early and Mid-Career Achievement in Research. He is Co-Editor of the *Journal of Policy Analysis and Management* and the *International Public Management Journal*. Dr. Potoski's research investigates public management and policy in domestic and international contexts, including public sector contracting and service delivery, environmental policy, and voluntary regulations. He is co-author with Aseem Prakash of the *Voluntary Environmentalists* (Cambridge, 2006) and Co-Editor of *Voluntary Programs: A Club Theory Approach* (MIT). Dr. Potoski received a PhD in Political Science from Indiana University in 1998 and a Bachelor's degree in Government from Franklin and Marshall College.

Dr. Matthew Potoski
Public Policy and Administration
Department of Political Science
Iowa State University
potoski@iastate.edu



Acknowledgements

This research is supported by the Acquisition Research Program in the Graduate School of Business and Public Policy, Naval Postgraduate School. We also wish to thank Dr. Keith Snider for his efforts on behalf of the Acquisition Research Program.

Special thanks to MITRE for extending their support and time, and for providing some useful information on prediction markets.

We wish to thank Adam Siegel of Inkling Markets for his suggestions on prediction market design.

Abstract

This paper discusses the applicability of prediction markets in Defense Acquisition projects, specifically in estimating their cost and schedule.

Several temporal and political factors can sometimes limit the effectiveness of traditional methods of project tracking and cost estimation, which may be overcome by using prediction markets. A prediction market provides an environment for traders to buy and sell contracts whose values are tied to uncertain future events. Efficient prediction markets have been shown to outperform available polls and other forecasting mechanisms.

There are various prediction markets based on different models and algorithms. Our focus is not to analyze these models, but to identify the design principles of implementing a proven prediction market model in a defense acquisition project. Some pilot studies have been carried out that provide insight into the behavior of the market participants. We found increased involvement of participants and greater interest in the projects to be major benefits. The areas that need to be considered in the design and implementation of markets are related to the participants (like, which traders to include); the information to be collected, or the stocks; the marketplace to be used; and the incentive structure to keep the participants motivated to trade.

Introduction

Historically, defense programs for weapons acquisition have taken longer, cost more and delivered fewer quantities and capabilities than planned (GAO, 2006).

Some of the reasons for such discrepancies include:

- Lack of discipline in estimation. Sometimes all the required factors or experienced people are not consulted during the estimation process. There is also a lack of transparency and accountability (GAO, 2006).
- Unrealistic expectations. The estimates can sometimes be too optimistic (GAO, 2008).
- Initial estimates tend to “anchor” expectations (Aranda & Easterbrook, 2005). Initial figures of estimates set too high expectations from the start, which is difficult to meet later on. Also, it is later on difficult to incorporate any modifications.

A recent study by RAND involved an analysis of data in the Selected Acquisition Reports (SARs) for a sample of 68 completed programs. The results showed that the average cost growth (including both cost overruns and cost under runs) for these programs



was 46% over the baseline estimate made at Milestone B and 16% over the baseline estimate at Milestone C. While cost growth occurred earlier in acquisition projects, the development cost growth at completion for programs initiated in 1970s, 1980s and 1990s remained relatively steady (Arena, Leonard, Murray & Younossi, 2007).

It is also worthwhile to note that of the 23 programs assessed in the GAO report (GAO, 2006), around 10 are expecting development cost overruns greater than 30%, or have delayed the delivery of initial operational capability by at least 1 year. The impact of such conditions is the reduction in the value of the DoD's defense dollars and lower return on investment (ROI) (GAO, 2006). See Table 4 in the Appendix for more details.

The key recommendation for the executive action from the report is to ensure that appropriate knowledge is captured and used at critical junctures, specifically at the following key points:

- Program start,
- Design review for transitioning from system integration to system demonstration, and
- Production commitment.

An effective way to capture knowledge and utilize collective intelligence of all team members is to use prediction markets or information aggregation markets. Such markets can be used as a supplement to existing estimation methodologies in generating more accurate cost and schedule estimates.

Prediction Markets: History and Purpose

Prediction markets are emerging as a promising forecasting mechanism for efficiently handling the dynamic aggregation of dispersed information among various agents (Tziralis & Tatsiopoulos, 2007). Various terms such as “information markets,” “decision markets,” “electronic markets,” and “virtual markets” can be used to describe this. In essence, prediction markets are speculative markets created for the purpose of predicting the outcome of an uncertain future event. They provide an environment for traders to buy and sell contracts whose values are tied to uncertain future events.

In the late 80s, some academics at the University of Iowa's business school came up with idea of giving students hands-on experience in trading markets, such as stock and commodities. Instead of using play money, they created a real market in which anyone could bet modest sums on the outcome of future events—e.g., the next president of the United States. The Iowa Electronics Markets (IEM) is now a thriving nonprofit enterprise (Tetlock, 2006).

Figure 1 shows a sample prediction market that is commercially available:



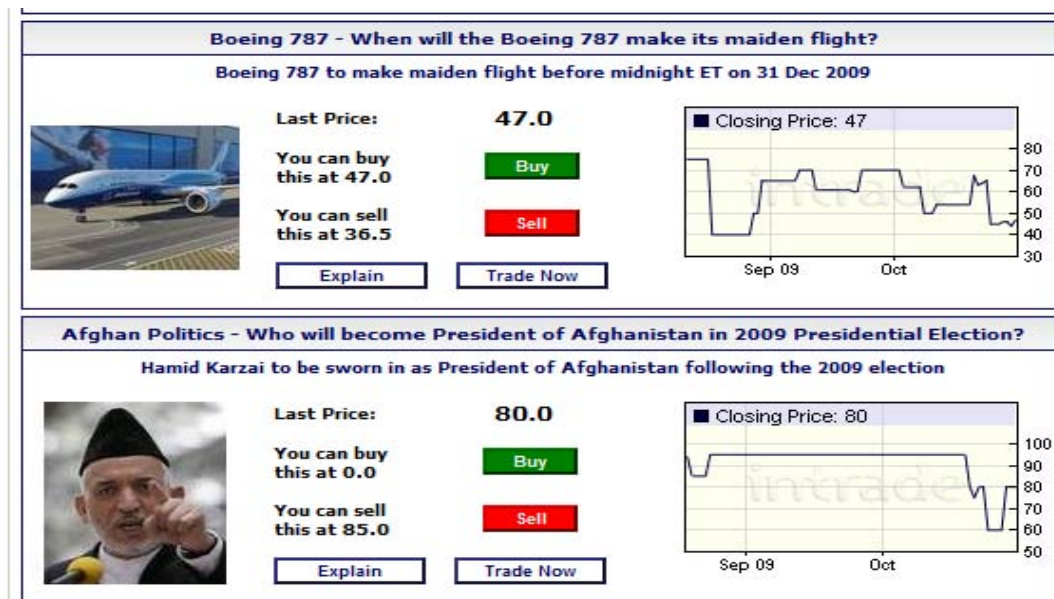


Figure 1. A Sample Prediction Market
(Intrade.com, 2010)

The theoretical foundations of prediction markets lie in the efficient market hypothesis, which states that a sufficient number of marginal traders with rational expectations, who maximize utility through maximizing profits, drive prices in the market in such a way that there is no opportunity for arbitrage (Shvarts & Green, 2007).

It is believed that if the market is well functioning, then contract prices reflect the collective wisdom of market participants. There are three primary types of prediction markets (Wolfers & Zitzewitz, 2004):

- **Winner Takes All Market:** In this, the contract pays a specified amount if and only if the world achieves some specific conditions, with the price of the contract reflecting the market's assessment of the probability of that occurrence.
- **Index Market:** In this, the contract pays an amount reflected in some condition of the world, such as the percentage of the vote a candidate receives or the inches of cumulative snowfall in a city.
- **Spread Market:** In this, contracts specify the cutoff that determines whether an event occurs (such as win margins in football). These basic contract types can be extended to generate additional predictions and measures of market uncertainty, such as the probability that a candidate will receive 40% of the vote, 50%, 60%, etc.

Table 1 illustrates some examples for each:

Table 1. Types of Prediction Markets
(adapted from Wolfers & Zitzewitz, 2004)

<i>Contract</i>	<i>Example</i>	<i>Details</i>	<i>Reveals market expectation of . . .</i>
Winner-take-all	Event y : Al Gore wins the popular vote.	Contract costs $\$p$. Pays $\$1$ if and only if event y occurs. Bid according to value of $\$p$.	Probability that event y occurs, $p(y)$.
Index	Contract pays $\$1$ for every percentage point of the popular vote won by Al Gore.	Contract pays $\$y$.	Mean value of outcome y : $E[y]$.
Spread	Contract pays even money if Gore wins more than y^* % of the popular vote.	Contract costs $\$1$. Pays $\$2$ if $y > y^*$. Pays $\$0$ otherwise. Bid according to the value of y^* .	Median value of y .

Prediction markets are most notably used today for predicting election outcomes, movie box office returns, terrorist attacks and sporting events (Wolfers & Zitzewitz, 2004). Information markets have also been deployed at Hewlett-Packard Corporation for making sales forecasts and were found to be more effective than traditional methods. “Not only did the market predictions consistently beat the official HP forecasts; the outcomes predicted are consistent with the probabilistic predictions of the IAM (Information Aggregation Market)” (Plott & Chen, 2002).

Google also launched an internal prediction market in April 2005. The Google markets are similar to Iowa Electronic Markets, and have survey questions like “How many users will Gmail have?” Common type of markets included those forecasting demand (Masse, 2008). “Google’s prediction markets are reasonably efficient, but did exhibit four specific biases: an overpricing of favorites, short aversion, optimism, and an under-pricing of extreme outcomes. New employees and inexperienced traders appear to suffer more from these biases, and as market participants gained experience over the course of our sample period, the biases become less pronounced” (Masse, 2008).

Some of the factors that are involved in prediction markets are:

- **Illiquidity:** Prediction requires enough buyers and sellers making enough transactions to produce a clear market price that summarizes the market prediction. In “thin” markets where buyer, sellers and their trades are few, the prediction market can sometimes be ineffective. Some research suggests prediction markets can be effective with as few as sixteen active traders (Christiansen, 2007).
- **Rationality:** While prediction markets have been strikingly accurate when they meet the criteria for effective markets, scholars have noted some deviations from “perfect” rationality. First, while traders’ preferences may bias their trades, such as when partisans buy political futures contracts for their favored candidate (Forsythe, Rietz & Ross, 1999), the prediction market price of contracts will remain accurate so long as the enough marginal traders remain objective (e.g., their profit motives outweigh their preferences). Prediction markets do not require all traders be rational as long as marginal market exchange is motivated

by rational traders (Gruca & Berg, 2007; Tziralis & Tatsiopoulos, 2007). Second, as in horse racing, prediction markets participants exhibit over price long shots (so that contract holders receives a smaller payoff relative to the true probability of the event occurring) and under price high probability events, though Cowgill, Wolfers, and Zitzewitz (2009) suggest that precisely the opposite biases occurred when Google employees traded in internal company prediction markets forecasting Google's future. Moreover, prediction markets may be less effective at accurately predicting extremely likely or unlikely events (Wolfers & Zitzewitz, 2004).

- **Manipulation and Bubbles:** A sufficiently endowed and motivated trader might attempt to manipulate a prediction market by purchasing contracts in desired directions. A candidate for elected office, for example, might purchase contracts on himself to generate apparent momentum and publicity. Speculative bubbles are possible in prediction markets, a non-trivial concern in light of recent macroeconomic events. Wolfers and Zitzewitz (2004) propose that a speculative bubble occurred in political prediction markets on whether Hillary Clinton would win the 2004 Democratic presidential primary, and bubbles may have occurred in experimental prediction markets reported in Plott and Sunder (1988). To some extent, markets may self correct manipulation and bubbles (Strumpf, 2004). The best defense against manipulation is sufficient liquidity so that profit seeking traders, recognizing that the manipulated market prices are inaccurate trade for contracts to bring predictions back into line.

Prediction Markets for Defense Acquisition Projects

The DoD's ability to meet its acquisition targets is becoming increasingly critical as defense budgets are reduced and expectations to deliver on time and on budget remain high. Costs overruns are one aspect of this issue. Please refer to Table 5 in the Appendix for a sample of cost overruns in large federal projects.

A recent GAO report (Edwards, 2009) concluded that "weapon programs are taking longer, costing more and delivering fewer capabilities than originally planned." It also noted that "systematic problems both at the strategic and at the program level" (GAO, 2008) were to blame. The GAO also noted that military branches "overpromise capabilities and underestimate costs to capture the funding needed to start and sustain development programs" (GAO, 2008).

In 2008, the GAO reviewed cost and schedule of 72 weapons programs and found that the average cost overrun for systems development was 40% (GAO, 2008) and concluded that "DOD's acquisition outcomes appear increasingly suboptimal" (GAO, 2008). A study by Deloitte consulting also agrees that defense cost overruns are getting worse (Irwin, 2009).

Table 2 gives a brief comparison between existing cost estimation techniques and the prediction markets. The goal here is to compare the probable shortcomings of the former with the possible advantages of the latter.



Table 2. Value Propositions of Prediction Markets

Probable limitations of existing cost methods	Possible value propositions of prediction markets
<ul style="list-style-type: none"> Cost Estimating Relationships (CERs) based on <u>technical factors</u> rather than programmatic “soft” factors. 	<ul style="list-style-type: none"> Shift of focus from estimating by <u>individuals to groups</u>.
<ul style="list-style-type: none"> <u>Not dynamically updated</u> as a program evolves, making the original estimate outdated as soon as the climate changes. 	<ul style="list-style-type: none"> Provide way to leverage information from <u>diverse sources</u>.
<ul style="list-style-type: none"> Manifestation of a <u>few decision-makers</u>, under tremendous time pressure, working with limited and perhaps biased information. 	<ul style="list-style-type: none"> <u>Mitigate biases</u> stemming from pressure to “price to win” or hide damaging information.
	<ul style="list-style-type: none"> Enable <u>frequent sampling</u> of information.
	<ul style="list-style-type: none"> <u>Incentivize traders</u> to seek out quality information that will help them do better.

Prediction markets are a judgmental forecasting method (Graefe, 2009) in which the price mechanism of the market automatically aggregates all the dispersed information among participants. Various aspects of the prediction markets are also discussed in further sections.

Design Principles of Prediction Markets

As with any markets, prediction markets may fail—and produce inaccurate forecasts—if not properly designed and executed. They should follow specific mechanisms, contracts and contexts to ensure participants are adequately incentivized to participate.

Some other factors to be incorporated are (Suda, 2009):

- Should require unbiased tacit knowledge of crowd,
- Should be designed in well-incentivized environments,
- Should be designed for a non-hierarchical environment,
- Generally not suitable for idea generation process,
- Should have well-defined end dates or closing criteria, and
- Choices should be mutually exclusive and have a definitive outcome.

Pilot Studies on Prediction Markets

In order to study the prediction markets, we conducted some simulations on a group of students at MIT to study the effectiveness of prediction markets, and some interesting observations were noted. Details of the study are given below.

Pilot Study 1: Predicting Ratings for Sunday Night Football

A group of 17 participants (N=17) were asked to participate in a private marketplace for predicting the Nielsen Rating (NBC – USA) for Sunday Night Football (Oct 25). Please refer to Figures 8, 9, and 10 in the Appendix for the screenshots of the market. The following rules were set in place:



- Time duration for trading was 5 days, with the market set to close 30 minutes prior to game start.
- Each participant was allotted an initial amount of \$5,000 for trading (play money).
- Winning Criteria was generating maximum worth.
- Incentive—Points for class participation grade and a popular book for winner.
- Minimum of four trades per person (tied to the incentive of class grade), with no maximum limit on trading.
- 11 rating ranges were presented as choices or contracts for participants to trade on, with each rating range being assigned equal probability (9.09% each) to start with (See Figure 2).
- Participants could “buy” or “sell” shares of the choice or contract they found most likely to occur. Every time a user bought a share for a particular idea, the price went up. Similarly, each time users sold a share, the price went down. The user’s account was also credited or debited based on his/her choice.
- There was a direct correlation between price and probability—share price of \$9.09 indicated 9.09% chance of that particular rating range.
- The buy and sell transactions were managed by an automatic market maker.

Results

Around 236 trades were conducted in total by 17 participants.

The highest rating range predicted was of (9.6-10) with a 19% likelihood prediction, followed by range (9.1-9.5) with a 17.3% chance (see Figure 3).

What will be the Nielsen Ratings for the NBC Sunday Football Game (Oct25)?

Question will be judged on: 10/25/09 @ 05:00 PM PDT

TIP: Current value = probability prediction will occur. e.g. \$10 = 10% chance prediction will occur.

Select a prediction:

PREDICTIONS	CURRENT VALUE	TODAY
Between 9.6-10.0	\$19.08	(closed)
Between 9.1-9.5	\$17.28	(closed)
Between 10.1-10.5	\$12.69	(closed)
Between 8.0-8.5	\$11.57	(closed)
Between 11.1-12.0	\$11.44	(closed)
Between 10.6 - 11.0	\$9.90	(closed)
Between 8.6-9.0	\$6.58	(closed)
Between 12.0 - 13.0	\$3.65	(closed)
Below 8.0	\$2.85	(closed)
Over 16.0	\$2.48	(closed)
Between 13.0 - 16.0	\$2.43	(closed)

Figure 2. The Stocks in the Prediction Market
(Inklingmarkets.com)

The actual rating range turned out to be (9.1–9.5).

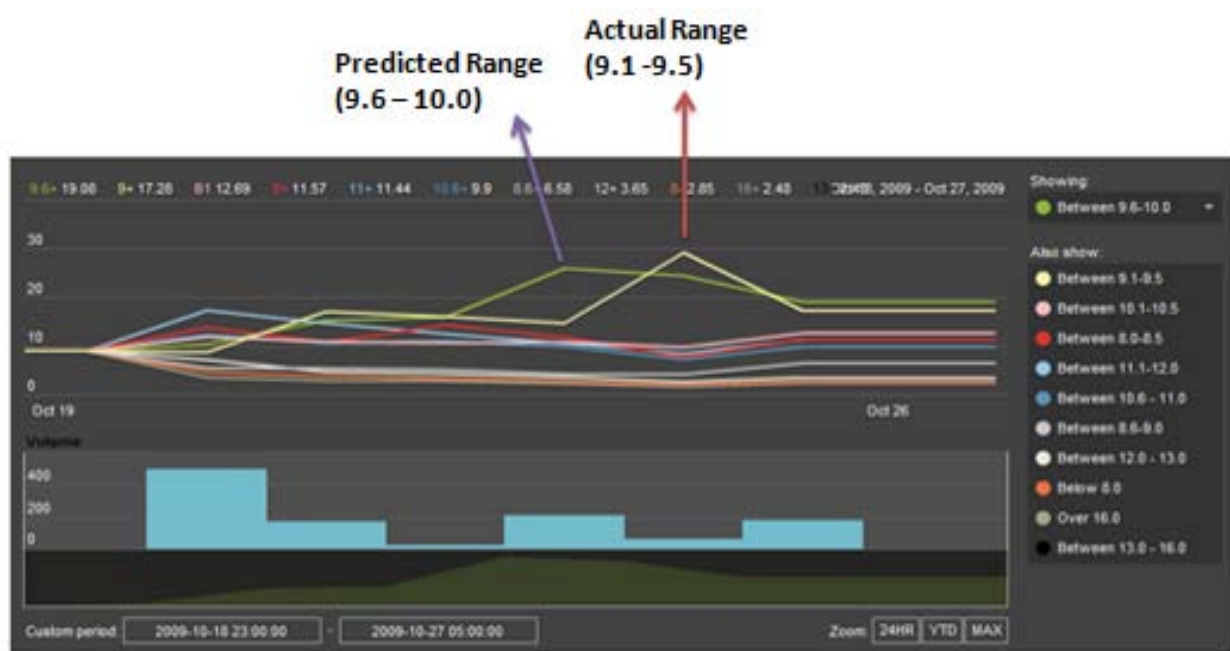


Figure 3. Results at the Close of Prediction Market

A probability distribution function (Pdf) for the result is shown in Figure 4, which demonstrates how close the predicted result was to the actual result.

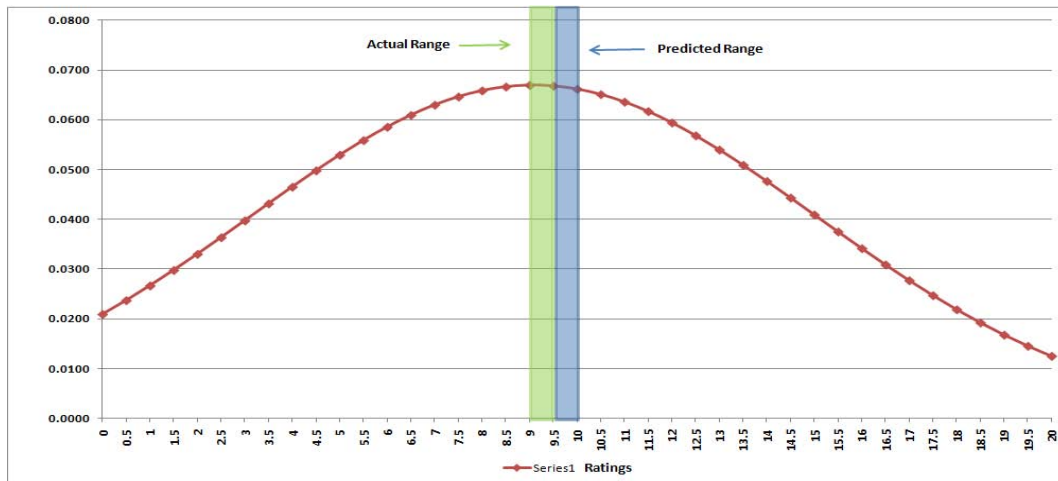


Figure 4. Pdf of the Prediction Market Results

Although market predicted range was close, there can be two reasons why the market forecast was not absolutely accurate:

- On the day of the game, a World Series game ran into overtime—overlapping with the football game and impacting the expected ratings for the Sunday night football. It turns out that there were no baseball fans in the participating traders, and hence this information was not captured.
- In a study on markets on NFL football games from tradesports.com, there was an observation of mispricing (Solomon & Hartzmark, 2008) visible primarily in

sporting markets. This was due to a behavioral bias known as disposition effect, which is the tendency of investors to sell stocks that have risen in value and hold on to stocks that have fallen in value, with the objective of earning maximum profit. Thus there might be mispricing in such scenarios and the hypothesis that the prices in prediction markets reveal unbiased estimates of true probabilities of the event happening might not be true.

Pilot Study 2: Predicting Ratings for Monday Night Football

Based on the previous pilot study feedback, the structure of the next market was modified. The participants (N=13) were asked to predict the Nielsen Rating (Cable TV - US) for Monday Night Football on ESPN. Instead of providing range choices to participants for bidding, they were asked to bid on an indexed market. Rules were similar to the last market, except for the following changes:

- Instead of ranges, participants now saw only one rating. They had to predict if the rating as per their opinion was to be higher or lower. Based on their responses, the price or rating was adjusted automatically.
- The initial rating set was 10.
- The price and rating scale was 1:1, i.e., a price of \$10.00 meant the current projected value of rating was 10.00 (see Figure 5).

Results

Around 116 trades were conducted by 13 participants. The final rating predicted by the market was 10.00 (see Figure 6).

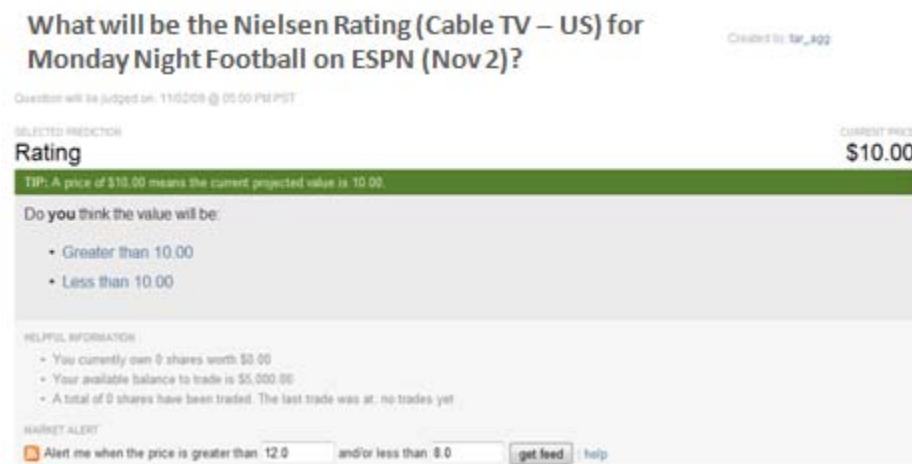


Figure 5. The Stocks in the Prediction Market
(Inkingmarkets.com)

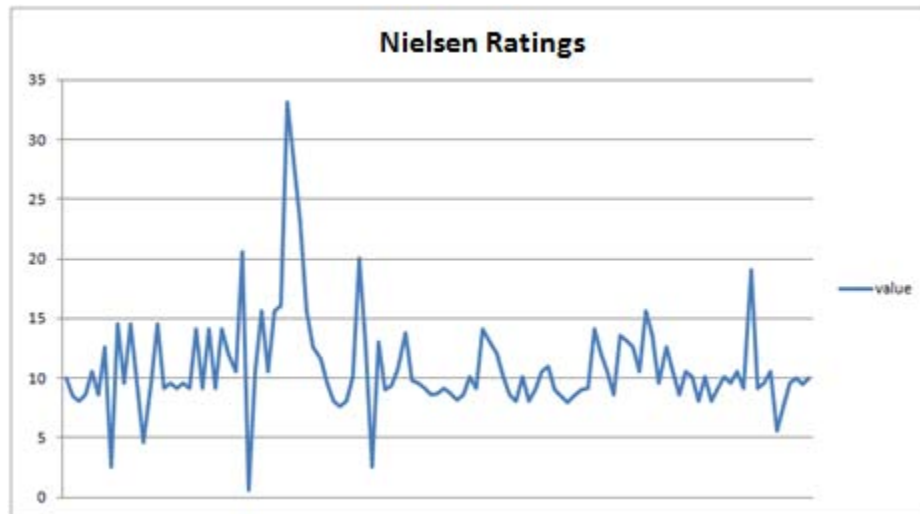


Figure 6. Graph Showing Change in Value of the Stock upon Trading

There was a huge variation in the ratings predicted, with the highest value being 33.1 and the lowest value being 0.6.

The actual rating turned out to be 7.7, which was different from the market value. The major reason for this was that market place was too volatile. A single participant could buy a large number of shares and alter the rating to a great extent, which is not a good design principle for prediction market.

Observations

Behavioral Observations of Participants

The participants comprised a mix of people with football knowledge and others with no knowledge at all. The actual result of the rating was quite close to the rating range that the test market predicted. But the main benefit was observed in the behavior of the participants. An unbiased, incentivized market piqued their interest and motivated them to remain involved throughout the market duration. Some other observations were:

- Increased participation due to internal competition,
- Attempts at market gaming,
- Trying to make use of First mover advantages,
- Forming alliances to manipulate the market,
- High incentive to participate led participants to gain more knowledge on the subject, and
- Use of statistical models by some participants to forecast ratings.

Other Useful Insights from Pilot Studies for Implementing Prediction Markets

- Ensure enough liquidity. Research done on prediction markets (Christiansen, 2007) suggests at least 16 participants for an effective market. Based on the two pilot studies, the number of participants was lesser in the second study, potentially explaining the volatility and resulting significant error in the prediction.

- Avoid volatility by choosing the right scale for trading in market. The market needs to be less volatile, so that it is not easily influenced by a single trader.
- Ease of use. The prediction market should be easy to use, as many participants were initially reluctant to trade due to perceived complexity.
- Ensure regular information update. In both the studies, there were not enough “happenings” regarding the game ratings, during the trading duration to sustain participants’ interest. The market is expected to fare even better if some new knowledge is available throughout the trading duration.
- Leads to increased awareness and interest. A significant contribution of prediction markets in the field of cost or schedule estimation is to keep the participants involved and make the process exciting. Traders with no knowledge of football got involved by researching more on the subject.
- Align incentives with the participants’ interests. Some people just executed the minimum required four trades and stopped. Although there was high trading, some people did not care to do all the research. One reason for this was due to incentives not aligning directly with the objectives of these participants.

Designing Prediction Markets for Defense Acquisition Projects

There are various components of a market that need to be considered prior to designing and implementing a prediction market for a project. A prediction market is composed of the following:

- **Stocks:** These are the outcomes or possibilities of the market that are collectively exhaustive and mutually exclusive. These act as stocks and securities for the traders to trade against, with a “buy” indicating a belief of occurrence of that particular outcome and a “sell” indicating a belief of non-occurrence.
- **Marketplace:** This refers to the software environment that contains the market or the questions for trading. It also encompasses the market manager that facilitates trades among the traders and manages the information distribution process.
- **Traders:** These are the participants or team members in a project that will be trading information in the form of stocks in the market.

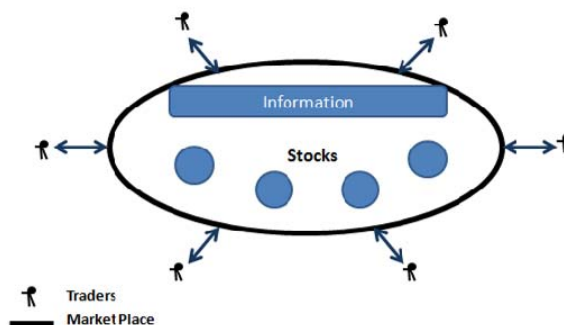


Figure 7. The Composition of the Prediction Market

There needs to be a continuous flow of information for the traders to base their decisions on and trade throughout the market duration.

Incentive structure is another important factor that should be considered during market design, as it acts as the motivating factor for the traders.

The success of the project hinges on identifying an acquisition program willing to participate. The ideal characteristics of the program are:

- uncertain cost and schedule components that can be specified clearly in a contract,
- a sufficiently large program community to ensure a liquid market,
- ample “soft” relative to “hard” information and information is broadly and unevenly held by diverse actors,
- one-of-a-kind program or a program with limited relevant historical information, and
- susceptible to performance impacts as a result of external events (i.e., political landscape, policy changes, personnel attrition, technology maturity, design modifications).

Based on initial pre-screening, we identified “Shallow Water Combat Submersible” (SWCS) as a suitable project, with a plan to implement the prediction market for duration of 3-6 months.

Below is some of the discussion on the design components for the prediction market.

1. **Stocks:** The stocks or the outcomes traded in the prediction market form the most important component of the market. The stocks traded will give us an indication of the beliefs of the participants, and the number of stocks traded will show us the level of confidence in their beliefs.

Hence, the questions that are asked in the market are very important.

There are mainly two categories of questions that can be asked.

- Type I questions: Asking these questions can directly provide the required information. $Q_1 \rightarrow A$
- Type II questions: Asking these questions can indirectly provide the required information, and a series of questions might be required for getting the information directly. $Q_2 \rightarrow Q' \rightarrow Q'' \rightarrow A$

This is similar to another perspective, as per which the state of the world can be divided on basis of two dimensions (Healy, Ledyard, Linardi & Lowery, 2009). Some questions can be based on an unobservable factor whose value can impact realization in second dimension, e.g., the underlying monetary policy of a central bank. While some questions can be based on a directly observable factor, e.g., interest rate each quarter.

The latter type of questions correspond to Type I questions, while the former compare to Type II questions.

Handling Type II Questions

The distinction of questions based on these categories is important because these allow us to focus more carefully on the “stocks” or outcomes of Type II questions, so that the required information is available.



- The suitable approach with the Type II questions is to ask multiple questions in the first round of prediction market, and, subsequently, drill down to specific questions in further rounds.
- If we consider a single question as a single market, then one of the approaches can be to run multiple markets at a given time, and then to add follow-up questions based on the responses. New markets (i.e., questions) can be run every week, and old questions can be updated weekly based on new information. The important factor with this approach is to decide which questions need to be deployed first.

Designing Questions for the Market

There is a lot of information that can prove useful in case of a defense acquisition project. This information can be either related to any factors impacting the cost or schedule of the project, or can be related to the current or future decision processes within the project. Based on the type of information required, we have further categorized the market questions.

- Questions of the First Order: These are the questions that give us the information about the program of interest. Some examples of questions for the pilot project are:
 - Will SWCS be certified by August 1, 2010?
 - The cost of the first unit will be \$x.
 - The final cost of the program will exceed the baseline estimate by x%.
- Questions of the Second Order: These are the questions that give us the information about traders and the trading process. For the pilot project, some of the questions are:
 - What is your motivation for trading (e.g., to win/to solve the problem/to validate my knowledge)?
 - What is your role in the Organization?
 - What are useful sources of information about the project?
- Questions of the Third Order: These questions give us information of the behavior outside of the prediction markets. Some of the questions are:
 - How have you used the prediction market in your job? If you notice a change in a contract price, how have you used that information?
 - Did trading improve your confidence in your opinion?

The First Order questions will be asked in the prediction market for trading. The Second Order and Third Order questions are more of profiling questions, which can help decide how to improve the markets, or how the participants are gaining from the market information. These questions can be sent to the participants as a survey and can be tied to incentives so as to encourage them to participate. Incentive can be additional bonus money (like, Earn \$100 more (of play money) for answering 5 questions).



The above factors helped to shortlist some of the questions to be used in the prediction market for the pilot defense project (SWCS). A risk mapping, showing the risk covered by a specific question of the First Order, along with the modified version of the question, is shown below. Note that the first four questions are of Type I, while the last question is of Type II.

Table 3. Risk Mapping for First Order Questions

Question Type	First Order Questions (Information about the program of interest)	Measured Risk	Alternate way of asking the question	
Type I	Will SWCS be certified by August 1, 2010?	Schedule Overrun	When will SWCS be certified?	Or
			By how many days will the certification date be exceeded?	
Type I	The cost of the first unit will be \$x.	Cost Overrun	What do you think is the most significant cost driver for the first Unit? Follow up question: Based on the below ranking of cost drivers for the first Unit, what do you think the cost of first Unit will be?	
Type I	The final cost of the program will exceed the baseline estimate by x%.	Cost Overrun		
Type I	What will be the 2011 Fiscal Year appropriation for the SWCS?	Cost and Schedule		
Type II	The SWCS program will be adequately resourced to meet its milestones?	Resource Allocation Errors	By how many days/months will the milestone be missed?	Or
			Which resource do you think is most necessary to meet the program's schedule objective?	Or
			Given the available resources, will the project be able to meet its planned schedule?	

Another important aspect is to mix some fun questions (like, who will win the Monday night game, etc.) along with serious questions to keep the participants involved.

- **Marketplace:** There are two important aspects of design for the marketplace—the software environment in which the participants perform trades and the market manager who facilitates the trading process.

Considerations for the Software Environment

The following are the design considerations for the software environment, based on observed behaviors in pilot simulations and requirements for a defense environment:

- The software environment will be easy to access and use.
- The software environment will be commercially available, with prior success in defense related environment.
- The environment will satisfy the IT security considerations—like secure login, user access rights, data security, etc.
- The software environment will have dashboard and administrative capabilities to analyze the data and generate reports based on the user behavior.
- It will be possible to share new information with the traders and/or close or open new markets whenever required.

Considerations for the Market Manager

The matching mechanism in prediction markets has most often been a continuous double auction in which computer software matches buyers offering bids with sellers and their asking prices. Prediction markets have been successful using real-world money, purely “play money,” with no economic value (beyond the satisfaction of “winning”), and “prize money” that can be exchanged for prizes and entries in prize drawings. We will establish prediction markets with play and prize money and will establish contracts for the cost and duration of particular features of the acquisition program. This is more realistic than focusing on the entire program completion, given the limited time we have to run the market. Fully specified, clear and enforceable contracts in prediction markets require the future state of the world (or events) on which the contracts can be easily adjudicated.

Market Scoring Rule, invented by Professor Robin Hanson at George Mason University, will be used in the prediction market. This will ensure that fewer traders can be present in the market, and lot of concepts related to stock markets that are difficult to understand (like bid-ask spreads, etc.) can be avoided (Siegel, 2010). The price of the stock represents whatever the last trader was willing to pay for shares and is set automatically according to the volume and sentiment in the question. For example, if someone buys 10 shares of an answer, the price of the share will be increased automatically according to an algorithm.

To ensure sufficient liquidity, we will use a prediction market with automated market maker, which means that the buying and selling of stocks will happen between traders and the market and not between the traders themselves.

- **Traders:** While prediction markets require some disagreement among potential traders over forecast, (else no trades would occur) excessive information heterogeneity can be harmful to the markets functioning. If some traders possess significant private information (and hence become insiders), the outsiders may refuse to participate, ultimately killing the market. Prediction markets are most likely to be successful if traders hold sufficiently balanced information about the event but have differing beliefs about the meaning of common information. It is also critical that when new information becomes available, it will be widely known.

We will recruit market participants from those employed in the Acquisition program or those surrounding the Acquisition program (i.e., subcontractors, end users, consultants). The traders in the defense project comprise of the different team members across different cross-functional areas. There will be many people in a project who will have limited information as to what is going on with a program at large, and they might be biased to their individual projects. Hence, it will be ensured that participants from all smaller projects, as well as people from other departments are included in the trading process. The acquisition project will have sufficiently broad following to attract participants from inside and outside the government. These can include administrative personnel, sub-contractors, engineers, end users, consultants, etc. The markets will be implemented for duration of 3 to 6 months, giving ample trading time to the participants.

Sometimes, due to fear of hierarchy or work environment pressures, people do not reveal uncomfortable information related to the project. In the prediction market, anonymity of traders will be maintained so that they can disclose useful information without hesitation.



- **Incentives:** The incentive structure is perhaps the most important component of prediction market in the DoD environment. Incentives should be such that they are non-biased and encourage participation independent of any hierarchical considerations. “As with any business incentive system, a considerable challenge exists in choosing incentives that motivate the right behavior. . . With information markets, incentives must serve a dual role: to motivate participation and to motivate participants to provide truth-revealing opinions. Incentives that satisfy both criteria can be difficult to define” (LaComb, Barnett & Pan, 2007).

Performance-based incentives in the project can be provided in various ways. For example, participants can be rewarded based on the portfolio value at the closure of the market. Awards can also be given to the most active traders, or the traders contributing the most valuable information.

Non-monetary awards can be as effective as monetary awards when used as incentives. For example, Google found that in their internal markets, participants cared more about non-cash prizes like t-shirts rather than cash prizes (Coles, Lakhani & McAfee, 2007). Similarly, by announcing user rankings, the play-money markets in Christiansen’s field experiment (2007) were successful, even without monetary incentives.

Other factors like social competition, recognition and opportunity to contribute towards the project can also act as important motivating factors.

For the pilot defense acquisition project, we feel the following incentives will encourage participation:

- Declaring winners based on the maximum portfolio value at market closure,
- Maintaining a leadership board displaying the top players, and
- Including results of the prediction market in regular status reports to track any new or less visible information.

Conclusion

In his book *The Wisdom of Crowds*, James Surowiecki (2005) argued that if we take a crowd of diverse people and correctly aggregate their judgments, we will be able to get more accurate results based on the collective intelligence of the crowd. This wisdom of crowds can be seen in action every day (Coles, Lakhani & McAfee, 2007) with the collective intelligence also driving Wall Street—the probabilities generated from the market displayed through the stock prices.

The prediction markets can be implemented for defense acquisition projects to aggregate the dispersed information across different functional areas. This information can help keep track of any cost or schedule variances, and also identify any potential risks that can impact the project. We anticipate prediction markets can outperform existing Defense Acquisition estimation techniques (i.e., parametric, analogy, activity-based (Boehm, Valerdi, Lane & Brown, 2005) in cases where there is ample “soft” relative to “hard” information and information is broadly and unevenly held by diverse actors.

Markets need to be compatible with the political and regulatory contexts in which they operate. For example, federal regulations permit play and prize money prediction markets but have restricted real money markets. Perhaps the most notorious prediction market was the ill-fated DARPA sponsored policy prediction markets (Hanson, 2007), which collapsed from a political firestorm over the (quite illogical) proposition that terrorists might benefit from trading in terrorism futures.



Based on the various case studies and literature available on prediction markets, our observations through simulations of markets, and discussion on various aspects of defense acquisition projects, we consider the following design principles:

- The acquisition project should have a sufficiently broad following to attract participants from inside and outside the government. These can include sub-contractors, end users, consultants, etc.
- Although play money should be used for participation in the markets, the traders can be incentivized based on their constant participation and ability to predict accurate results.
- Ample historical data should be available on similar projects, so that participants who are new to the project have some basis of gaining knowledge and predicting.
- The questions for the market should be very carefully designed so that the required information can be revealed. “Fun” questions should be mixed with “hard” questions so that the participants remain involved.
- To ensure sufficient liquidity, prediction markets with automated market makers should be used, and at least 15-16 participants should be targeted for trading (Christiansen, 2007).
- Results of the prediction markets should be carefully studied and reported as part of regular status reports.
- Anonymity of participants with respect to responses should be maintained to encourage information-sharing.
- The length of the market and closure of each specific question need to be given proper consideration while designing the market. The support of senior management in Defense projects is essential for starting the prediction market.
- Markets need to be compatible with the political and regulatory contexts in which they operate.

As part of the ongoing research project, the observations and feedback from preliminary and literature research will be used to run the field trial with the DoD (Department of Defense) acquisition programs for a fixed duration (up to 6 months). The prediction market for the selected pilot acquisition project still needs to be implemented, and other design and implementation factors will be added or the current ones revised based on new information.

If prediction markets improve in forecasts in the defense arena, as they do in other venues, our aim will be to assist in implementing and then studying additional markets, including those over the longer term and a wider variety of programs. It is clear from our research that a significant contribution of prediction markets in the field of cost or schedule estimation is to keep the participants involved and make the process exciting. Used with right implementation guidelines, these markets can prove to be helpful in capturing the collective wisdom of the project members for the specific duration.

Appendix

Table 4. Development Cost Overruns by Decade
(GAO, 2006)

1970 - 1979	1980 - 1989	1990 - 1999
Development cost overrun: \$13 billion (30%)	Development cost overrun: \$12 billion (39%)	Development cost overrun: \$15 billion (40%)
Key Studies and Initiatives Impacting the Defense Acquisition Process		
<ul style="list-style-type: none"> • 1970 Fitzhugh Commission • 1972 Commission on Government Procurement 	<ul style="list-style-type: none"> • 1981 Carlucci Initiatives • 1982 Grace Commission • 1986 Packard Commission 	<ul style="list-style-type: none"> • 1994 Federal Acquisition Streamlining Act • 1996 Clinger-Cohen Act
DOD Acquisition Policy Changes		
<ul style="list-style-type: none"> • 1971 DOD 5000 policy established • 1975 Policy revised • 1977 Policy revised 	<ul style="list-style-type: none"> • 1980 Policy revised • 1982 Policy revised • 1985 Policy revised • 1986 Policy revised • 1987 Policy revised 	<ul style="list-style-type: none"> • 1991 Policy revised • 1996 Policy revised

Source: DOD (data); GAO (analysis and presentation).

Table 5. Sampling of Federal Cost Overruns
(Edwards, 2009)

<i>Defense Development Costs (\$2008)</i> ⁶³	<i>Original Estimates</i>	<i>Final Estimates</i>
Global Hawk surveillance plane	\$989m (2001)	\$3.7b (2007)
Expeditionary Fighting Vehicle	\$1.6b (2000)	\$3.6b (2007)
C-130J Hercules	10.9m (1996)	430.3m (2007)
Extended Range Munitions	86.9m (1997)	500.1m (2007)
DDG 1000 destroyer	2.2b (1998)	9.3b (2006)
V-22 Osprey helicopter	4.0b (1986)	12.5b (2006)
Armed Reconnaissance Helicopter	388.3m (2005)	750.9 (2007)
Space Based Infrared System High	4.2b (1996)	8.5b (2006)
Armed Reconnaissance Helicopter	388.3m (2005)	750.9 (2007)
NPOESS Satellite System	5.0b (2002)	7.9b (2007)
Other Defense		
Coastal Patrol Ships ⁶⁴	\$220m (2004)	\$350m (2007)
Joint Strike Fighter ⁶⁵	\$232b (2001)	\$337b (2008)
Marine One (VH-71) helicopters ⁶⁶	\$6.1b (2005)	\$11.2b (2008)
Coast Guard, NSC ships, per unit ⁶⁷	\$250m (2002)	\$536m (2007)

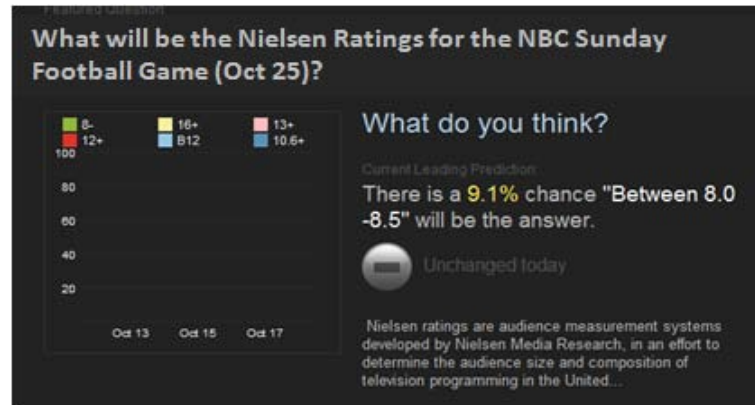


Figure 8. Screenshot 1 from the Prediction Market Simulations (inklingmarkets.com)

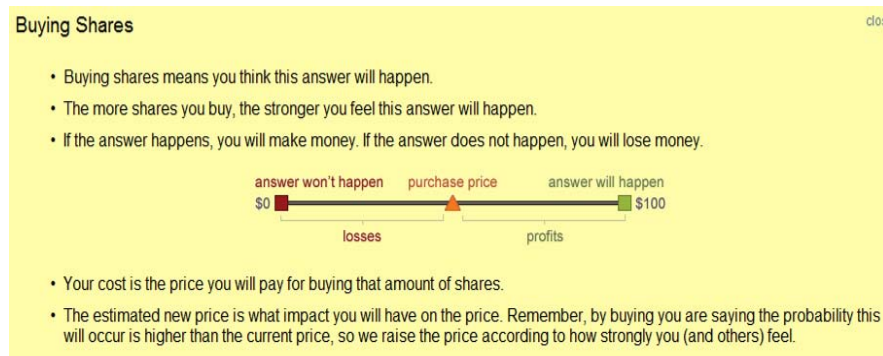


Figure 9. Screenshot 2 from the Prediction Market Simulations (inklingmarkets.com)

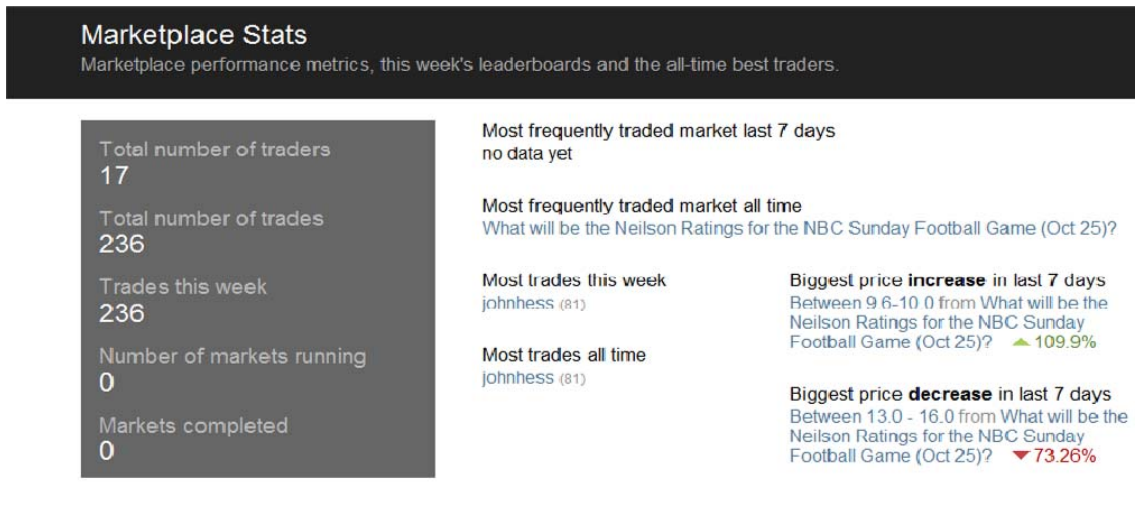


Figure 10. Screenshot 3 from the Prediction Market Simulations (inklingmarkets.com)

List of References

- Aranda, J., & Easterbrook, S. (2005). *Anchoring and adjustment in software estimation*. European Software Engineering Conference on the Foundations of Software Engineering.
- Arena, M.V., Leonard, R.S., Murray, S.E., & Younossi, O. (2007). Historical cost growth of completed weapon system programs (TR-343-AF, 2006). Santa Monica, CA: RAND Corporation. Retrieved September 19, 2007, http://www.rand.org/pubs/technical_reports/TR343/
- Boehm, B.W., Valerdi, R., Lane, J. & Brown, A.W. (2005). COCOMO suite methodology and evolution. *CrossTalk - The Journal of Defense Software Engineering*, 18(4), 20-25.
- Christiansen, J. D. (2007). Prediction markets: Practical experiments in small markets and behaviors observed. *Journal of Prediction Markets*, 1(1), 17–41.
- Coles, P., Lakhani, K., & McAfee, A. (2007). *Prediction markets at Google* (HBS case study).
- Cowgill, B., Wolfers, J., & Zitzewitz, E. (2009). *Using prediction markets to track information flows: Evidence from Google* (Working paper). Philadelphia, PN: University of Pennsylvania.
- Edwards, C. (2009). *Government-cost-overruns*. Retrieved December 20, 2009, from downsizinggovernment.org: <http://www.downsizinggovernment.org/government-cost-overruns>
- Forsythe, R., Rietz, T., & Ross, T. (1999). Wishes, expectations and actions: Price formation in election stock markets. *Journal of Economic Behavior and Organization*, 39(1), 83–110.
- Galton, F. (1907). Vox Populi. *Nature*, 75, 450-451.
- GAO. (2004, May). Lack of disciplined cost-estimating procedures hinders effective program management (GAO-04-642). Washington, DC: Author.
- GAO. (2006, April). *Defense acquisitions: Major weapon systems continue to experience cost and schedule problems under DOD's revised policy*. Retrieved November 9, 2009, from <http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=gao&docid=f:d06368.pdf>
- GAO. (2006, November). *DoD needs to take more action to address unrealistic initial cost estimates of space systems* (GAO-07-96). Washington, DC: Author.
- GAO. (2008, March). *Defense acquisitions: Assessments of selected weapon programs* (GAO-08-467SP). Washington, DC: Author, p. 4.
- GAO. (2008, June 3). *Defense acquisitions: Better weapon program outcomes require discipline, accountability, and fundamental changes in the acquisition environment* (GAO-08-782T). Washington, DC: Author.
- Graefe, A. (2009, March 25). Prediction markets versus alternative methods: Empirical tests of accuracy and acceptability. Dipl.-Volkswirt, Dipl.-Wirtsch.Inf.Andreas Graefe.
- Gruca, T.S., & Berg, J. E. (2007). Public information bias and prediction market accuracy. *Journal of Prediction Markets*, 1(3), 219-231.



- Hahn, R.W., & Tetlock, P.C. (Eds.). (2006). *Information markets: A new way of making decisions*. Washington, DC: AEI-Brookings Press.
- Hanson, R. (2007). The policy analysis market—A thwarted experiment in the use of prediction markets for public policy. *Innovations: Technology, Governance, Globalization*, 2, 73-88. MIT Press.
- Healy, P.J., Ledyard, J.O., Linardi, S., & Lowery, J.R. (2009). *Prediction market alternatives for complex environments*. Columbus, OH: Ohio State University.
- Irwin, S. (2009, January 1). Weapons cost overruns: From bad to worse. *National Defense*.
- Intrade.com. (2010). Screenshot. Retrieved March 2, 2010, from www.intrade.com
- Home.inklingmarkets.com. (2010). Screenshot. Retrieved March 20, 2010, from <http://home.inklingmarkets.com>
- LaComb, C.A., Barnett, J.A., & Pan, Q. (2007, July). The imagination market. *Information System Frontiers*, 9(2-3), 254.
- Masse, C. (2008). Do Google's enterprise prediction markets work?—Using prediction markets to track information flows: Evidence from Google. Retrieved March 25, 2010, from Midas Oracle .ORG: <http://www.midasoracle.org/2008/03/26/google-enterprise-prediction-markets-4/>
- Plott, C. & Chen, K.-Y. (2002, March). *Information aggregation mechanisms: Concept, design and implementation for a sales forecasting problem* (Working Paper No. 1131). Pasadena, CA: California Institute of Technology Social Science.
- Plott, C.R., & Sunder, S. (1988). Rational expectations and the aggregation of diverse information in a laboratory security markets. *Econometrica*, 56, 1085–1118.
- Siegel, A. (2010). Retrieved March 2, 2010, from http://inkingmarkets.com/homes/frequently_asked_questions
- Shvarts, R., & Green, S. (2007). Cost-benefit analysis and prediction markets and the proposal for University of Oregon prediction markets.
- Solomon, D., & Hartzmark, S. (2008, May). Efficiency and disposition effect in NFL prediction markets.
- Strumpf, K. (2004). *Manipulating the Iowa political stock market*. Mimeo, University of North Carolina.
- Suda, B. (2009, August 14). *Introduction to prediction markets*. Retrieved March 1, 2010, from <http://optional.is/required/2009/08/14/prediction-markets/>
- Surowiecki, J. (2005). *The wisdom of crowds*. New York: Anchor Books.
- Tziralis, G., & Tatsiopoulos, I. (2007). Prediction markets: An extended literature review. *Journal of Prediction Markets*. Retrieved January 28, 2010.
- Wolfers, J., & Zitzewitz, E. (2004, May). Prediction markets. *The Journal of Economic Perspectives*, 18(2), 107-126.
- Wolfers, J., & Eric, Z. (2005). *Interpreting prediction market prices as probabilities* (Working Paper No. 10359). Cambridge, MA: National Bureau of Economic Research.

Innovations in Defense Acquisition: Implementing Information Aggregation Markets

Peter Coughlan—Peter J. Coughlan is an Associate Professor of Business and Economics at the Graduate School of Business and Public Policy, Naval Postgraduate School. He received his PhD in Economics from the California Institute of Technology in June 1999. Prior to joining the Naval Postgraduate School in 2004, Dr. Coughlan was an Assistant Professor of Business Administration in the Strategy Group at the Harvard Business School. Dr. Coughlan's research interests include experimental economics and game theory/mechanism design. He also has research interests in strategic and competitive dynamics, particularly applied to technological change.

Peter J. Coughlan
Graduate School of Business and Public Policy
Naval Postgraduate School
Monterey, CA 93943-5000
E-mail: pjcoughl@nps.edu

Jeremy A. Arkes—Jeremy A. Arkes is an Associate Professor of Economics in the Graduate School of Business and Public Policy at the Naval Postgraduate School. He received his PhD in Economics from University of Wisconsin. His other professional positions have been as a Research Analyst at the Center for Naval Analyses and an Economist at the Rand Corporation. He has published articles and reports on teenage risky behaviors, military manpower, PTSD, and military logistics.

Jeremy A. Arkes
Graduate School of Business and Public Policy
Naval Postgraduate School
Monterey, CA 93943-5000
Phone: (831) 656-2646
E-mail: jaarkes@nps.edu

William R. Gates—William R. Gates is an Associate Professor of Economics and Dean of the Graduate School of Business and Public Policy at the Naval Postgraduate School. He received his PhD in Economics from Yale University in 1984. Before joining the Naval Postgraduate School in 1988, Dr. Gates worked as an economist at the Jet Propulsion Laboratory in Pasadena, CA. Dr. Gates' research interests include game theory (matching games and auctions), incentives and asymmetric information (incentive contracting) and public goods (economics of defense alliances).

William R. Gates
Graduate School of Business and Public Policy
Naval Postgraduate School
Monterey, CA 93943-5000
Phone: (831) 656-2754
E-mail: bgates@nps.edu

Abstract

Prediction markets are information aggregation mechanisms that have received increasing attention in the civilian sector. Prediction markets, sometimes called information aggregation markets, idea markets or event futures, are essentially small-scale electronic markets that tie payoffs to measurable future events. They are similar to stock markets, where the "stocks" are outcomes or events rather than shares in a company. This research will address prediction market design and implementation issues and provide



recommendations on how prediction markets might be used to address acquisition-related issues such as program cost, schedule or performance estimates, prospects for successfully completing project milestones or meeting performance criteria in system tests, etc. (although these securities clearly need to be better defined).

Research Issue—The maxim that “the devil is in the details” is especially true for designing and implementing such markets. In particular, significant research, testing, and calibration must be conducted prior to execution to be confident that the forecasts produced by any prediction market are (1) actually forecasting the measure intended and (2) providing the most accurate forecasts achievable via this methodology. This research will use a literature survey and a pilot prediction market test and experimental analysis to address prediction market design in an acquisition environment.

Research Result—Preliminary results regarding prediction market design issues. This research will be in-going throughout FY2010, but preliminary results will be available.

Keywords: Prediction Markets, Information Aggregation Markets, Idea Markets, Event Futures



THIS PAGE INTENTIONALLY LEFT BLANK



Panel #16 – Considerations for Managing Service Oriented Architecture Projects

Thursday, May 13, 2010

9:30 a.m. –
11:00 a.m.

Chair: Reuben Pitts III, President, Lyceum Consulting, LLC

An Exploration of the AT&L AV SOA System

Robert Flowe, Office of the Deputy Under Secretary of Defense (Acquisition & Technology) and Maureen Brown, The University of North Carolina at Charlotte

Meaningful Cost-Benefit Analysis for Service-Oriented Architecture Projects

Lloyd Brodsky, CSC

Service-Oriented Architectures and Project Optimization for a Special Cost Management Problem Creating Optimal Synergies for Informed Change between Qualitative and Quantitative Strategic Management Processes

Stefan Pickl, Goran Mihelcic, and Marco ShulerThe University of Bundeswehr Munich



An Exploration of the AT&L AV SOA System

Robert Flowe—Robert M. Flowe is the Enterprise Integration Functional Team Lead for the Acquisition Visibility Service Oriented Architecture (AV SOA) initiative, within the Directorate of Acquisition Resource Analysis (ARA) within the Office of the Under Secretary of Defense Acquisition Technology and Logistics (OUSD(AT&L)). In this role, Flowe guides the elicitation, definition, and implementation of functional/analytical requirements for acquisition data visibility. Acquisition data visibility ensures complete, timely, and accurate program data is made visible to authorized consumers to support effective analysis, oversight and decision-making within AT&L and the broader defense acquisition workforce. Flowe's role in modernizing acquisition data visibility through technologies such as service-oriented architectures, Web services, and value-added business intelligence tools, reduces the cost and effort associated with data production and reporting, while providing greater insight into the status and behaviors of the defense acquisition enterprise, thereby enabling better understanding and informed decision-making at all levels of acquisition management and oversight.

Prior to his current position, Flowe led the Strategic Studies program within the Systems Analysis branch of the Directorate for Systems Engineering within the Office of the Director, Defense Research and Engineering. In this role, he established and managed the research agenda for the Systems Engineering Research Center (SERC), a University-Affiliated Research Center dedicated to system engineering research.

Previously, Flowe served as a senior operations research analyst within the Office of the Secretary of Defense (OSD) Cost Analysis Improvement Group (CAIG). During his seven years on the CAIG staff, he led and supported independent cost analyses on a variety of major defense acquisition programs, including submarines, space launch vehicles, command, control, communications, computers and intelligence (C4I), and automated information systems. Flowe retired from the United States Air Force in 2003, after having served in a variety of acquisition, engineering and operational positions relating to space launch; intelligence and information operations, specializing in systems and software acquisition. While on active duty, he served on the faculty of Defense Acquisition University, where he managed and taught the Intermediate Software Acquisition Management Course. Flowe has a BS in Aerospace Engineering from Virginia Tech, and an MS in Software Systems Management from the Air Force Institute of Technology, and is Level III Certified in Program Management, and Program Systems Engineering.

Robert M. Flowe
AV SOA Functional Team Lead
Office of the Under Secretary of Defense Acquisition Technology and Logistics Enterprise Integration
OUSD(AT&L)/ARA/EI
1225 S. Clark St., Suite 900
Arlington, VA 22202
Phone: (703) 699-0183
Blackberry: (571) 309-6506

Abstract

This proposal reports the findings of an NPS-funded research effort on the goals, objectives, and properties of the AV SOA system. Every academician and graduate student knows all too well that scholarly research is often limited by the availability of data. Without data, hypotheses remain untested, theories remain unproven, and understanding is thwarted. Nowhere is this more true than the DoD acquisition enterprise. Institutional incentives and fractured priorities have undermined efforts to gather timely, valid, complete, and authoritative data on defense acquisition activities. The resulting isolated, incomplete, disaggregated, hidden, missing and just plain wrong data results in acquisition policies



formed by anecdotal evidence, untested across the broader DoD experience base. Management theories developed in the 1960s are assumed to be equally valid in the 21st century. Even when objective evidence suggests something may be profoundly wrong, the lack of data and corresponding research obscures the path forward.

To address this crisis, AT&L ARA, first with the Defense Acquisition Management Information retrieval (DAMIR) system and, more recently, with the Acquisition Visibility Service-Oriented Architecture (AV SOA) and other initiatives, seeks to make authoritative, current, and complete acquisition data more visible to the community than ever before.

This paper describes the suite of data, business intelligence tools, and other products now available to the acquisition research community. The authors will outline the research objectives and analytical methods that may be developed through the use of these resources and provide examples of the striking insights these data can provide when used in thoughtful research.

Keywords: AV SOA authoritative data defense acquisition



Meaningful Cost-Benefit Analysis for Service-Oriented Architecture Projects

Lloyd Brodsky—Lloyd Brodsky is a solution architect senior consultant in the Federal Consulting Practice of Computer Sciences Corporation. He frequently consults on governance and acquisition issues in service-oriented architecture for a variety of government clients. He holds a doctorate in management information systems from the Sloan School of the Massachusetts Institute of Technology and an MBA in finance from the business school of the University of California, Berkeley. He also holds the PMP and CISSP certifications and completed COTR training while employed at the Department of Veterans Affairs.

Abstract

This paper argues that proper cost-benefit analysis of service-oriented architecture projects is not possible without explicit identification of SOA-specific tasks in the work breakdown structure (WBS), so that those costs are explicitly estimated in the budget, are explicitly in the integrated master schedule, and appear on earned value and other reports. It deconstructs the traditional stories for financially justifying SOA and identifies SOA-specific activities that should be added to WBSs to enable tracking of costs and schedules. It also identifies specific research questions that can only be answered with data gathered through such task-level cost accounting.

Introduction

The central point of this paper is quite simple: You cannot do proper return on investment (ROI) or earned value analysis (EVA) of a service-oriented architecture (SOA) project without a work breakdown structure that explicitly identifies SOA-specific activities in a way that facilitates comparing investments and benefits. This often does not happen because SOA has additional steps in development and operation that either do not exist or are less important in traditional development. This paper deconstructs the traditional arguments for SOA to identify these activities and shows how those costs come to be commingled with other development and maintenance activities. The paper argues that this comingling impairs the ability to diagnose problems and recommend best practices. It recommends that acquisition professionals require inclusion of important SOA tasks in the work breakdown structure and the integrated master schedule, as well as requiring reporting on a range of SOA cost issues.

Service-Oriented Architecture Review

SOA is a term with multiple meanings. It refers to both an architectural style and to web services technology. The architectural style calls for dividing needed functionality into separately built business function-related services. Applications are composed of such loosely coupled services, which usually communicate with each other via a web-service interface. The underlying idea is that it would be cheaper and faster to build or modify applications by composing them out of limited-purpose components that can communicate with each other because the components strictly adhere to interface rules. There are three standard arguments for why this would result in financial savings and be a positive net present value deal (+NPV):



- **Interfaces:** SOA replaces $m \times n$ point-to-point interfaces with $m+n$ service interfaces. In exchange for an upfront investment in defining a common vocabulary and common interfaces and a continuing cost in governance and ESB maintenance, this would save maintenance costs over time. The savings would come from:
 - Having fewer interfaces to maintain
 - Reduced “information archeology” costs when making changes due to the tighter configuration management needed to get services to work
 - Reduced certification and accreditation costs due to fewer and better documented interfaces and services
 - The hiding of connection details by the enterprise service bus, if one were used¹⁸.
- **Authoritative databases:** SOA stores data once and publishes it as a service for all the applications that need it. This reduces costs by:
 - Only storing one version and,
 - Indirectly through separating application code from the data and,
 - Indirectly through better data quality – if you have multiple versions either they are all the same or something is wrong.
- **Reuse:** Money is saved to the extent that new applications and changes are accomplished through reusing existing services. An example would be reusing a security service in all applications that required authentication before use.

Cost Accounting Issues

As managers, we would naturally be interested in projecting these costs when planning and in gathering actual data in production to see how accurate those estimates were. There are also a number of more narrow SOA cost-accounting issues we would like to drill down to. These include:

- **Governance overhead:** A central part of SOA doctrine is the absolute requirement of a centralized governance organization that creates local standards and enforces both those and other applicable standards to ensure interoperability and architectural conformance. This review layer adds cost and time, but what the appropriate share of budget should be is not well understood.
- **Vocabulary synchronization cost:** It is also SOA doctrine that legacy systems can be made available as services by building an interface layer that maps the existing vocabulary to the ontology. It is quite possible that a “market analysis” of the demand for potentially-sharable information could lead to savings from not sharing information that is in little demand.
- **Timing of ESB and registry installation:** Enterprise service busses and service registries play a valuable role in masking connection complexity and in making developed services findable. However, due to rapid technological evolution of

¹⁸ Applications that read from flat files need to be changed if the record size or layout changes, whether that change affects the program or not.



COTS offerings, there is some question as to the value of introducing their use before a critical mass of available services is achieved.

- **Certification & accreditation relationship:** The standardization and tighter configuration management associated with SOA should drive down C&A costs.
- **Software best practices support services:** It is tacitly assumed that web-services interfaces are both well documented and stable, so that third parties can successfully use them by following the instructions. While good documentation and configuration is considered a best practice whether doing SOA or not, the possible damage is greater in SOA deployments.

The Work Breakdown Structure Problem

A work breakdown structure (WBS) is a hierarchical decomposition (tree structure) of the work required to accomplish a goal. It should be developed by starting with the end objective and successively redividing it into manageable components in terms of size, duration, and responsibility. However, it is often done as a modification of an existing WBS for a similar project. It is a required part of the DoD acquisition process.

The WBS is an essential starting input to both estimation and to scheduling. In essence, it provides the chart of accounts for a project. To get to a schedule, each task will have a duration, labor hours, and predecessor and successor tasks assigned to it. Pricing that labor gives a budget, not including materials. Each task in the WBS has hours, labor in labor categories and a duration. All earned value reports, all scheduling and all progress payments are keyed to the WBS. All work has to belong to and be “billed” to one of the tasks defined in the WBS. It follows that if you want to know what something costs, it needs to exist as a task in the WBS.

In large projects, the WBS is quite complex and can be as much as five or six levels deep. Usually, items at the same level of hierarchy are in the order they are executed, although this is not required. Traditionally, definition of the WBS is left to vendors with the integrated master schedule and price proposal based upon it included as part of the RFP response. More often than not, the organization of the WBS in software development follows the traditional “waterfall” method of system development. The primary constraint is that the WBS fulfills the requirements of the statement of work. In the Defense context, the foundation for WBSs is in *DoD Directive 5000.1* and in *MIL-HBDK-8881A*. The latter became a military standard on 1/9/2009. The Project Management Body of Knowledge published by the Project Management Institute (the basis for the PMP exam) also emphasizes the importance of the WBS in project management.

Since the development of the software WITHIN services is about the same as traditional development, we suggest that the distinctive feature of SOA from a WBS perspective is the tasks associated with developing the interfaces BETWEEN individual services. Unfortunately the practice of using project managers without a background in enterprise architecture has led to the development of WBSs that look more like traditional “waterfall” development, which leaves implicit the governance and common interoperability infrastructure, which is independent of any single service’s development and support team. It follows that the most helpful approach is to:

- **Divide the work into interfaces and services:** In SOA development there is a whole series of activities, such as ontology and interface development, whose function is to enable communication between services. These SOA-specific activities should be separately identified in the WBS.



- **Explicitly account for non-SOA activities that are critical for SOA:** While such activities as configuration management and technical documentation are not SOA activities per se, they are so important to SOA success they should be separately trackable as well.
- **See that the operations WBS is consistent with the development WBS:** Operations is often done by a different contractor and/or a different solicitation. If the idea is that, say, investment in web-service interfaces in development pays off in reduced cost of maintenance and change later, it would be helpful if the two WBSs facilitated that comparison.
- **Define relevant reports:** Because the value chains implied by the SOA benefit stories are fairly complex, some creativity is needed to define meaningful reports that aggregate associated things. For example, the cost of making changes to a service pursuant to a change order will involve governance review, coding, testing, independent validation, and certification and accreditation. In a mature SOA environment, this will involve not only aggregating managerial across different services maintenance organizations; it is also likely to involve different contractors. Understanding the complete cost impact and aggregating the information may be a challenge.

While there are not hard and fast rules for constructing a WBS in information technology development, the most common approach is to have tasks at the same level of hierarchy appear in order of start. Thus, in a number of WBSs examined anecdotally for this paper, at the top level they started with Enterprise Architecture and ended with Post-Deployment, with Development having the deepest structure and the largest number of leaf nodes.

Enterprise Architecture and Development are the two top-level activities most affected by SOA. To be consistent with the separation suggested above, this paper suggests the following new activities:

- **Enterprise architecture:** Enterprise architecture (EA) includes a diffuse range of engineering planning activities, which are bunched at the beginning of development but continue throughout development and into operations. EA is responsible for governance—the establishment of standards and subsequent review for compliance, which is essential to SOA success. Specific SOA tasks that would fall under governance involve:
 - **Planning**
 - **Ontology:** the development of controlled vocabularies for data interchange.
 - **Interface standards:** The standards that XML schemas and other technical artifacts are to follow. These are needed to implement standardized interfaces and their associated management and error correction protocols.
 - **Configuration management practices:** If future development and maintenance are to be able to connect to a service based solely on the technical documentation, configuration management and change management practices need to be very well controlled. EA needs to review and publish the definition of what a service is supposed to do and the service level agreement it is expected to follow.

- **Development review**
 - **Governance review:** The governance entity will need to review the interoperability artifacts and service contracts produced by the development team.
- **Run-time review and enforcement:** The governance entity needs to review the monitoring of service-level agreements and take appropriate corrective action in case of violations. Governance should be added. Also, configuration management should be increased.
- **Development:** By far the largest time and cost in any project is actually doing the development. This is typically subdivided into:
 - **Requirements and Design:** Specification and design the technical artifacts to implement planned interfaces. In addition, if a service is to be done by wrapping an existing capability, the vocabulary in that legacy system needs to be mapped into the ontology standard.
 - **Code and Unit test:** Implements the SOA-specific activities designed above.
 - **Integration test:** Independent verification and validation.

Conclusion

There is an old saying that you cannot manage what you cannot measure. By increasing the number of “moving pieces” in IT solutions, SOA increases the number of pieces that require measurement. Given the relative immaturity of the SOA paradigm, it is particularly important now, when best practices have not yet been established and the understanding of cause and effect is limited. Indeed, the inability to collect cost and schedule data at the task level may be part of the reason why so many case studies in SOA only present project-level estimates of averted cost.

It is well within existing authority for acquisition to require the WBS to make explicit anything that acquisition and the PMO would like to monitor. This will, in turn, assure that those SOA tasks appear explicitly in the integrated master schedule and on EVM reports. It is also possible for acquisition to standardize the terminology of WBSs across contracts in the same program, which could help assure tying investments made in development to their hoped-for payoff in operation. We recommend this happen and that acquisition also require reporting on SOA cost-benefit across the system lifecycle.

Service-Oriented Architectures and Project Optimization for a Special Cost Management Problem Creating Synergies for Informed Change between Qualitative and Quantitative Strategic Management Processes

Stefan Pickl—Stefan Pickl studied mathematics, electrical engineering, and philosophy at TU Darmstadt and EPFL Lausanne 1987-93 (Dipl.-Ing. '93, and Doctorate 1998 with award). He was assistant Professor at Cologne University (Dr. habil. 2005) where he obtained the *venia legendi* for "Mathematics". He is visiting Professor at University of New Mexico (USA), University Graz (Austria), University of California at Berkeley and at Naval Postgraduate School Monterey (USA). He was visiting scientist at SANDIA, Los Alamos National Lab, Santa Fe Institute for Complex Systems and MIT. He is director of COMTESSA (Competence Center for Operations Research, Management of Intelligent Engineered Secure Systems and Algorithms) which he founded in 2007 at UBWM. The Center is associated with Centre for the Advanced Study of Algorithms (CASA, University of Nevada Las Vegas), with Center for Network Innovation and Experimentation (CENETIX, NPS Monterey) and with SEED Center for Datafarming (Simulation, Experiments & Efficient Designs). Stefan Pickl is vice-chair of EURO (European Society for Operations Research) group "Experimental Operational Analysis" and chair of the advisory board of the German Society for Operations Research. He is in the scientific advisory board of the German Society for Disaster and Emergency Management. He published more than 100 contributions and presented more than 200 talks. In '03, '05, '07 international best paper awards.

Stefan Pickl
University of the Federal Armed Forces of Germany
Institute for Theoretic Computer Science
Mathematics and Operations Research
Werner Heisenberg-Weg 39
85577 Neubiberg, Germany
Phone +49 89 6004 2400

Goran Mihelcic—Goran Mihelcic successfully graduated from the University of Cologne where he studied Business Informatics with a focus on Operations Research as well as Information Management and Organizational Theory. Since 2007, he has been working as a research assistant of Prof. Dr. Pickl at the Department of Operations Research, University of the Federal Armed Forces in Munich. IT-based decision support frames the center of attention for his doctoral thesis. His special interests lay in the field of developing effective and efficient methods and procedures in the domain of crises management and critical infrastructure protection. Methods and best practices in the field of business intelligence, business process management, computer supported collaborative work and service oriented architectures build the further focus of his work.

Goran Mihelcic
University of the Federal Armed Forces of Germany
Institute for Theoretic Computer Science
Mathematics and Operations Research
Werner Heisenberg-Weg 39
85577 Neubiberg, Germany
Phone +49 89 6004 2400

Marco Schuler—Marco Schuler is an active Officer of the Federal Armed Forces of Germany. During his military career he successfully graduated from the University of the Federal Armed Forces in Munich where he studied Computer Science with a focus on Operations Research as well as



Modeling and Simulation. Since 2007, he has been working as a research assistant of Prof. Dr. Pickl at the Department of Operations Research, University of the Federal Armed Forces in Munich.

Marco Schuler
University of the Federal Armed Forces of Germany
Institute for Theoretic Computer Science
Mathematics and Operations Research
Werner Heisenberg-Weg 39
85577 Neubiberg, Germany
Phone +49 89 6004 2400

Abstract

Mobility plays a central role in the support of military staff of the German Federal Armed Forces. This demand most often is fulfilled by a central organizational unit which allocates needed vehicles out of a local car pool of the military facility.

One essential maxim is to meet the “approved” demand for mobility for any military employee of that facility at any time.

This paper is based upon the experience out of an optimization project that has been conducted at a large German military facility with about 3000 employees.

The optimization effort aimed at two dimensions:

- Optimization at business process level (qualitative)
- Optimization at the cost level (quantitative)

A short introduction is given into the overall process from the application for a vehicle to the allocation of the needed vehicle. After analyzing the old process and its inefficiencies, a proposal for an improved process design supported by a service oriented software approach is given.

The second part of this paper is focused on potential mathematical optimization approaches that can be chosen to reduce cost and make “intelligent” allocations to the given demands.

The demanding goal was a user-friendly decision support system that is able to make intelligent allocations.

Description of Former Global Car Demand Process

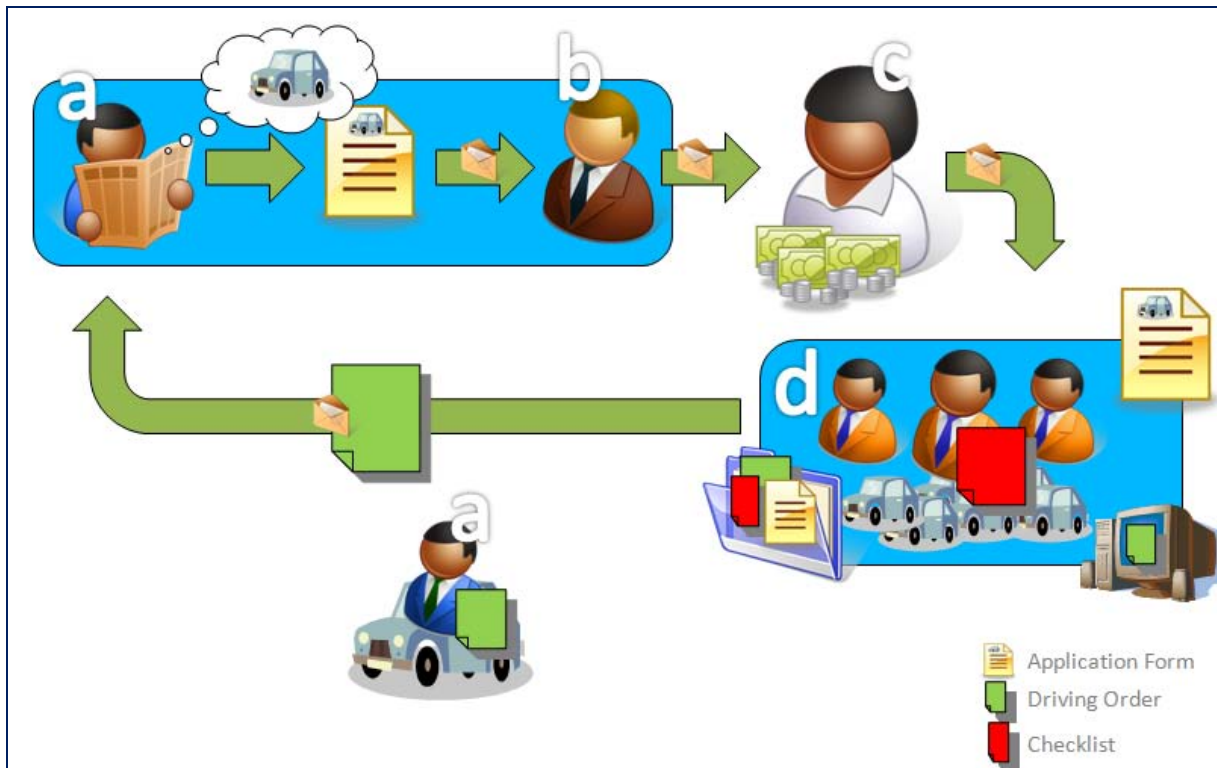


Figure 1. The Former Inefficient Process

The overall application process for a car starts with an application form (yellow document icon in Figure 1) that has to be filled out by the demanding person (a). Each department has an authorized person (b) who decides if the need for the car is appropriate and if the application can be granted. After having granted the application form, person (b) transfers the application form to the cost center (c) to ensure the funding of the car demand. If the funding is ensured the form is transferred to the local car pool management department (d). Here, the allocation of the car is conducted which includes the creation of a driving order (green document icon in Figure 1) and a checklist (red document icon in Figure 1).

The local car pool consists of a number of vehicles in short term and long term rent; if the request for a car cannot be satisfied out of the local car pool, (e.g., because of the unavailability of the same or a higher classed car), the request is escalated to a decentralized mobility centre which has access to a larger car fleet. Only if this mobility centre is not able to answer this demand either, an external car rental company is contacted to provide the desired car.

The driving order is the official document to certify that a certain typed vehicle has been assigned to a certain driver respectively group of drivers. This document is sent to the driver and confirms him officially to receive a car from the local car pool. On the day when the rental period starts, a person from (d) hands over the car to (a) and completes a checklist that contains information about the mileage, fuel tank level, potential damages, and additional equipment inside of the car. The driver's obligation is to document each (sub-) trip that he undertakes during the rental period; this information is captured on the driving order

form (green document icon, Figure 1) and contains information about the driver (important when a group of drivers are given), the number of kilometers, and the starting and end time of the trip. When the car is returned by the driver, a person from (c) checks the car for damages, mileage and fuel level and completes the checklist. The driving order which contains information about the trips made as well as the checklist remains with the local car pool department. The global process ends when the car is returned and the complete and consistent documents were handed in to the local car pool department.

One essential process step is the allocation of the car. Here, several parameters have to be taken into account in order to facilitate a cost-efficient and “intelligent” allocation that considers a balanced utilization of the car pool in terms of e.g. rental time and mileage. The task of “intelligent” allocation is fulfilled by a person of the car pool department (d) and is based on his experience in this field. The only it-based support so far consisted of an Excel-sheet that listed all the different types of vehicles with their corresponding performance-/statistical data (type, mileage ...) and a software tool which generated the driving order. When the car is returned, the Excel-sheet is updated with the trip data that has been collected by the driver on his driving order.

A more detailed reflection on the quantitative optimization methods that are needed for an intelligent disposition of a car as well as for the optimal mixture of rental types (short-/long term) is given in the second part of this paper.

Looking at the global process in Figure 1 in terms of process efficiency and usability, it becomes clear that this process is neither efficient nor user-friendly; documents are only handled manually, they are sent by internal post mail or they are faxed, (equal) data is gathered at several points of the process and very often redundantly, and the quality of data is very poor. All these factors lead to a very inefficient, time-consuming, and error-prone process.

Assuming that each postal service step (depicted by the brown envelopes in Figure 1) needs approximately two working days to reach its destination, the overall processing time consumes up to two weeks in order to provide the official driving order to the person requesting a car. This implies that urgent and short-term requests have nearly no chance to be granted in time so that induced by the given inefficiencies the requesting person might get into trouble to accomplish his task he needed the car for.

Description of Improved Global Car Demand Process

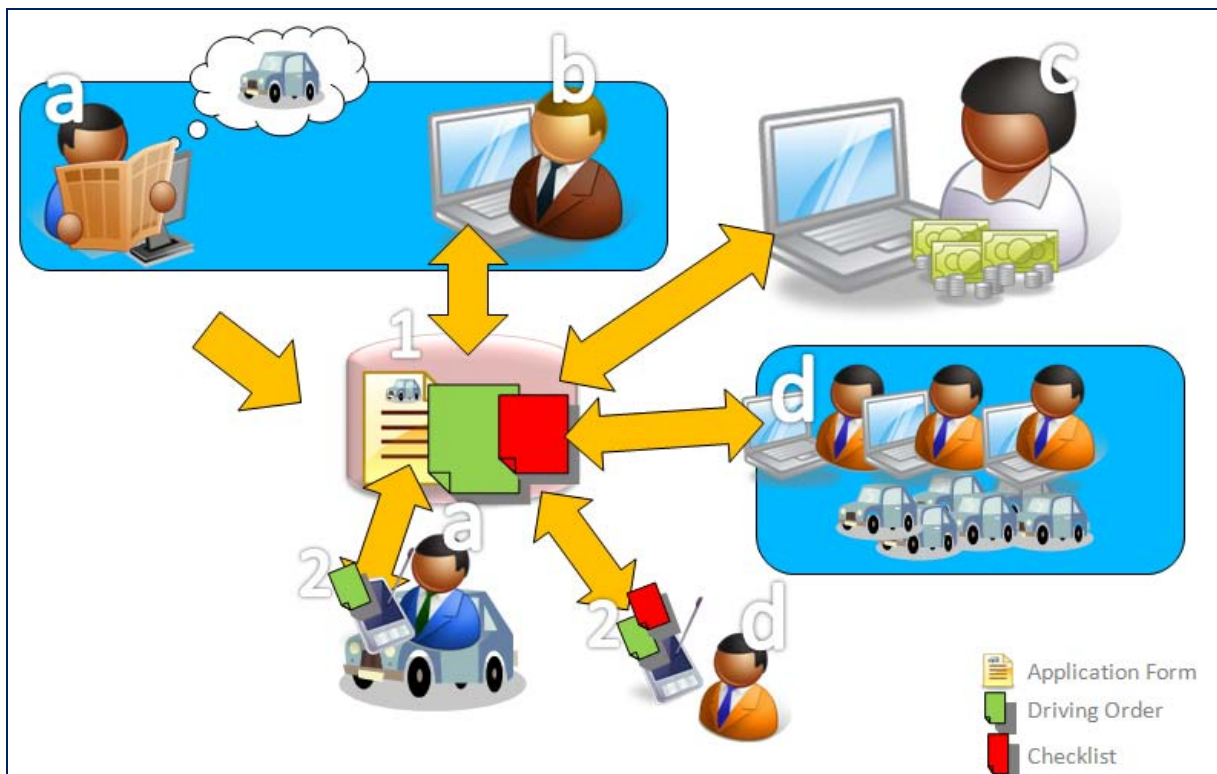


Figure 2. The Improved Process

In order to improve the process efficiency and the usability, it is indicated to introduce a software-based approach with a centralized data management and to define a workflow that supports the overall process from the application for the car until the return of it.

Central objectives should be a user-friendly, non-hardcopy document handling as well as a short processing time and an increased data quality in order to establish optimized decision making.

The introduction of a central database ("1" in Figure 2) that can be accessed by all organizational units that are involved into the global process, hard-copy documents become superfluous and the processing time can be minimized through the elimination of the postal document handling. The central data management minimizes redundant data collections which can now be stored in a controlled way into the database.

Another improvement, which leads to an increase of the data quality, is the introduction of a mobile device (as e.g., PDA or Smartphone) combined with an integrated gps-component and an optional GSM- or WiFi-module ("2" in Figure 2) for data transmission. This device supports the car pool service assistant in storing the checklist data into the central database. The device can then be handed over to the person that rented the car. For the trip-management, a special mobile application is running on the device and supports the driver in collecting the needed data. The driver-process that is support by the mobile application includes the authorization of the driver, indicating the start of the trip ("start-trip" button when starting trip), automatic registration of the number of kilometers driven, indication of the end of the trip ("end-trip" button at finish). When the car is returned, the device data is stored into the central database. Another option would be the real time

transmission of the data using the GSM-Module—but as this option might contradict to privacy policies it is not taken into further account.

A Flexible Service Oriented Architecture Approach

As it can be seen from the description of the old process, the legacy system implied many unnecessary steps and no databases or electronic forms were used.

An improvement can be reached in the development of a software solution given above that would use Service Oriented Architecture (SOA) concepts and principles. Additionally a SOA based approach would allow easily for further extensions and integration of complementary applications.

Figure 3 gives a more formalized presentation of the improved global process using the notion of event driven process chains (Scheer, 1998; Balzert, 2008).

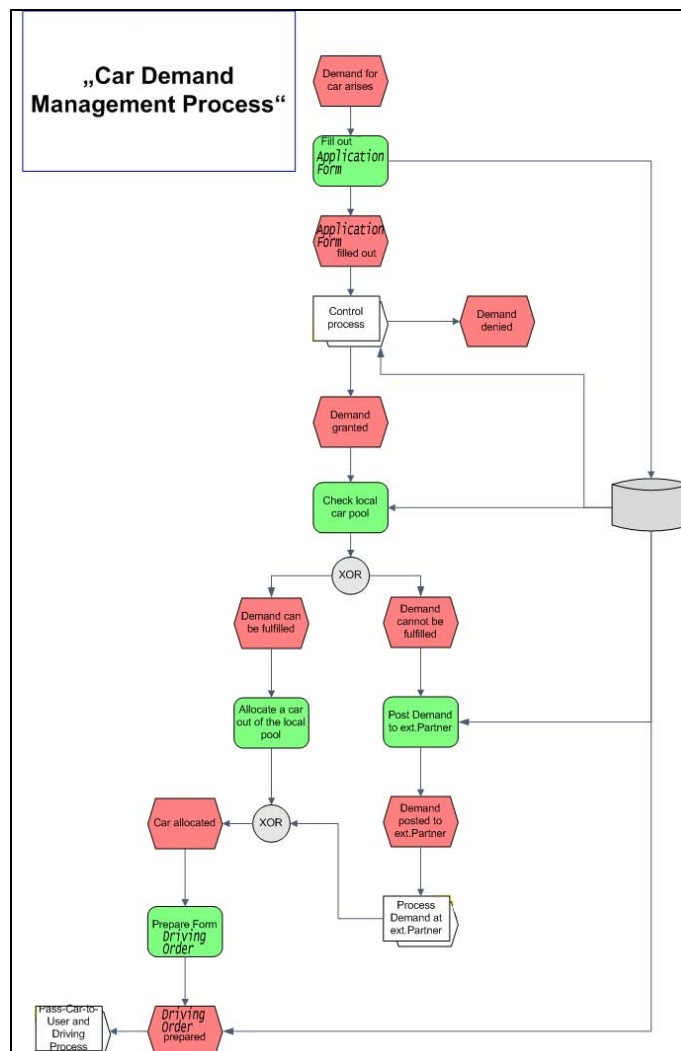


Figure 3: Event Driven Process Chain of Improved Process

Following the paradigm of matching business services to web services (WS)—assuming the use of web services as means of service oriented design patterns (see e.g., BPEL 2007; Drawehn 2008)—the following major services can be identified:

- Application Form WS
- Control Process WS
- Local Car Pool Allocation WS
- External Partner Allocation WS
- Driving Order Generation WS
- Car Handling WS

The Application Form WS handles the application form and supports the user in requesting a car by collecting all relevant data including consistency checks. etc.

The Control Process WS works on the application form data stored in the central database and supports the controller in handling the request.

The Local Car Pool Allocation WS is the “intelligent” component that is in charge of a cost efficient and balanced allocation of a car to the given request.

As mentioned above there may be situations in which the request cannot be satisfied out of the local car pool; in these cases, the External Partner Allocation WS is activated.

The Driving Order Generation WS generates the necessary driving order for the specific request.

When the car is picked up at the local car pool centre, the driving order is handed to the driver and the Car Handling WS is processed; this service initializes the mobile device with the needed data and manages the data handling of the checklist and trip data (at rental period start and finish).

A service oriented software approach enables an agile, efficient and user-friendly realization of the described “Car Demand Management Process” and it supports the flexible adaption of single services to the demand of the involved organizational units as well as the integration of external business partners (Krafzig, 2007).

Key aspect of the contribution is the connection of the service-oriented framework (“architecture”) with the classical optimization approach. As the problem describe an actual challenge within acquisition research an integral solution might support in an holistic and comfortable way the policy-making process for Department of Defense officials.

In the following we characterize the several optimization problems which are central. It might be observed that such approaches exist so far in several units but integration is missing and required. Furthermore, our contribution proposes an IT-based decision support tool which is characterized by an actual state of the art SOA-based framework.

In the following the different optimization “units” which should be integrated will be presented.

Description of the Problem

The overall optimization intent is to minimize the cost of the car fleet while several constraints have to be fulfilled. One constraint is that each driving job has to be served. As soon as the whole car fleet is busy an additional driving job creates the need of a new car.



While the base fleet is present anyway, thus producing no reducible cost, each additional car creates extra costs. The additional cars are the main source of cost of the present optimization problem. Nevertheless the inability of the existing fleet to cover all driving jobs might not be the only reason to go for more cars. There's another cost factor. All cars of the fleet are rent for a certain time frame—usually two years—and a certain mileage. If that mileage is exceeded an additional gradual charge is to be paid. Contrariwise an under-run mileage leads to a gradual rebate. That system could produce situations where an additional car would be the cheaper solution although the existing car fleet could cover the demand of driving jobs. As an additional difficulty, the driving jobs are categorized in certain classes of different types and qualities of cars.

In all models there are two main types of question to be answered. The first one is: "What car should execute what driving job?". Of course one car cannot deal with two overlapping jobs and the car has to be available in the fleet during all time frames of the driving jobs it runs. Second question to be answered is: "Are there better/cheaper solutions to the problem when more cars are added to the fleet?". The second question seems to be even more difficult to answer since there's no preset given for the number of cars to rent and rent time. It's also possible to rent an additional car in advance without even knowing the jobs it has to run. This adds the wide field of probability, prediction and uncertainty to the present problem.

Problem Categorization

To solve an optimization problem, it's often beneficial to categorize it. Finding the right problem category can be a tremendous step to solve the problem since with the problem category the algorithms and methods used to solve such a problem or the non-existence of such become apparent. The following sections will discuss possible categories and categorization factors. When modeling the present optimization problem there are several options selectable that will not only affect accuracy but also the complexity of the model. In some cases, it might be beneficial to reduce accuracy in favor of analyzability.

Linear versus Non-linear

One categorization factor is the linearity or non-linearity of the objective function on one side and of the constraints on the other side. For linear systems, there's in general a wider and better performing selection of algorithms and methods available to solve the problem. Therefore, linearity can be a decisive attribute of a problem. Unfortunately, the present optimization problem is apparently non-linear, since the distance constraints for the mileage of individual cars create non-linear gradual cost steps. Nevertheless, the main core of the problem can be modeled in a linear way omitting the mileage constraints. This leads to a less accurate but maybe better analyzable reduced model. Therefore the present problem is—depending on the analysis needs—either non-linear or linear. Thus for now no limitations are raised based on the linearity of the problem.

Integer versus Continuous

Most decision variables in the present problem are integers. Theoretically, the lease times for additional cars could be modeled as non-integers. But since the driving jobs are bound to a fixed length cycle a more flexible lease time for additional cars is of no use. Therefore, the present problem is an integer problem. Following that result, most of the



discussed problem categories in the following chapter originate from the combinatorial optimization.

Deterministic versus Probabilistic

Next to the already mentioned problem of assigning the right car of the fleet to the right driving job, including the option of adding new cars to the fleet, there's another optional addition to the problem. In order to improve the decisions made based on the model it might be beneficial to predict future driving jobs. This adds the whole complexity of randomness and decision under uncertainty to the present problem. As already explicated the added complexity might circumvent detailed enough analysis. Therefore, careful judgment is required whether driving job prediction is needed and useful.

Another kind of randomness comes into play when watching the mileage of the fleet cars. An incoming driving job comes with a very rough driving schedule that can be used to derive a prediction for the amount of kilometers the car will drive for the job. This prediction will again add a lot of complexity to the problem. But even without predicting the distance the car will drive, after its return the number of driven kilometers is a random parameter to the problem. Thus, as soon as mileage is included, the present problem becomes probabilistic.

Possible Problem and Solution Categories

Regarding the results from the previous chapter some promising candidates for the right problem category are discussed. Beside a short discussion of the solution strategy in general the pros and cons of the respective category in regard to the present problem are illustrated.

Exhaustive Enumeration

Omitting probability the most direct approach is exhaustive enumeration. This could be implemented as a simple back-tracking algorithm utilizing a depth-first search. Sadly the amount of possible combinations grows tremendously with the number of cars in the fleet and the number of driving jobs. In the worst case scenario the amount of combinations to check is

$$\sum_{k=0}^n \left[\frac{1}{k!} \cdot \sum_{i=0}^k \left[(-1)^i \cdot \binom{k}{i} \cdot (k-i)^n \right] \cdot (F+1)^k \right]$$

where n is the number of driving jobs and F is the number of cars in the Fleet. This grows terrible. A small example with only five cars and ten driving jobs already produces over 4.8 billion possible combinations to check. This is without additional cars and without predictions for future driving jobs in consideration. Without further improvements exhaustive enumeration is not feasible for real world scenarios.

Branch & Bound

One possible improvement to the exhaustive search is the branch and bound algorithm. It also tries to implement a depth-first search but it's looking for clever shortcuts to avoid as many branches of the search tree as possible. To achieve this, the algorithm tries to predict upper and lower bounds for all branches. With better upper/lower bounds the algorithm gains speed. Finding a good function for the upper/lower bounds is main

challenge of the algorithm. First attempts to utilize Branch & Bound for the present problem have shown that finding adequate bounds is very difficult. So far in most example situations the calculation time of the bounds exceeded the time advantage of the Branch & Bound algorithm compared to the exhaustive search. This is not because the calculation took so long but because the calculated bounds only removed very few branches from the search tree. Unless better functions for upper/lower bounds are found Branch & Bound is no improvement to the solution of the problem.

Scheduling Problem (Job Shop Scheduling)

It's possible to model the problem as a scheduling problem with the following attributes. The cars of the fleet translate into the machines of the scheduling problem. Machines of certain types can substitute certain other types, others can't. The driving jobs represent the tasks. The jobs have a fixed start and end time and therefore a fixed length. All jobs have a type or class defining on what type of machine they can be run. If no feasible schedule with existing machines can be found, new machines are added. Also, if feasible schedules can be found, there might be cheaper schedules with more machines taking mileage into account. That demonstrates the difference of the present problem to usual scheduling problems. A usual scheduling problem seeks the fastest completed feasible schedule while here the cheapest feasible schedule is to be found.

Summary and Outlook

Building flexible structures that combine qualitative measures of e.g. business process design with quantitative optimization approaches prove to provide very effective and efficient solutions in real-world problems of the daily military work. The Service-Oriented Architecture Approach serves as an excellent connector and enabler between these two "worlds." This approach offers a reasonable chance to be successfully transferable also to further domains of every-day military problem solving.

References

- Balzert, H. (2008). Lehrbuch der softwaretechnik. *Spektrum Verlag*.
- Web services business process execution language (bpel) version 2.0: Technical report*. (2007, April). OASIS.
- Drawehn J., Höß, O., Weisbecker, A. (2008, September). Prozessbasierte Anforderungsanalyse im Kontext von Serviceorientierten Architekturen (soa). In K.-P. Fähnrich, S. Kühne, & M. Thränert (Eds.), *Model-driven integration engineering* (pp. 51-60). Leipziger Beiträge zur Informatik (Vol. XI). Leipzig, Germany: Eigenverlag Leipziger Informatik-Verbund (LIV).
- Krafzig, D., Banke, K., & Slama, D. (2007). Enterprise SOA. Best Practices für Serviceorientierte Architekturen - Einführung, Umsetzung, Praxis, Mitp-Verlag.
- Scheer, A.-W. (1998). Wirtschaftsinformatik. Springer Verlag.

Panel #17 – Strengthening the Industrial Base

Thursday, May 13, 2010

**11:15 a.m. –
12:45 p.m.**

Chair: Dr. Michael McGrath, Vice President, Systems & Operations Analysis,
Analytic Services Inc.

Quadrennial Defense Review *Priorities and Implications for Acquisition*

Brett B. Lambert, Office of the Deputy Under Secretary of Defense
(Acquisition & Technology)

Requisites for a Strong Defense Industrial Base

Lawrence P. Farrell, Jr., Lt. Gen., USAF (Ret.), National Defense
Industrial Association

The Defense Budget and Defense Industry Finance

David Berteau, Roy Levy and Matthew Zlatnik, Center for Strategic &
International Studies



Quadrennial Defense Review Priorities and Implications for Acquisition

Brett B. Lambert— Brett B. Lambert joined the US Department of Defense (DoD) in 2009 as Director of Industrial Policy in the Office of the Deputy Under Secretary of Defense (Acquisition & Technology) (DUSD (A&T)). Mr. Lambert serves as the principle advisor to the USD (AT&L) on all matters relating to the defense industrial base, including industrial capabilities and assessments; defense industry mergers, acquisitions and consolidation; preservation of essential industries and technologies; and other related matters.

Prior to joining the DoD, Mr. Lambert was a managing director for the Civitas Group. In this role he provided strategic advisory services for companies operating in the intelligence, homeland defense, and national security sectors.

From 1989 until 2007, Mr. Lambert held positions of increasing responsibility at DFI International, a national security consultancy he built with the founder. Mr. Lambert assisted in the sale of DFI in 2007.

At the conclusion of his tenure with DFI, Mr. Lambert held the titles of Executive Vice President of DFI International and Managing Director of DFI Investment Partners. He was also the company's ethics officer and served as a member of the Board of Directors. Mr. Lambert specialized in technology assessments, strategic planning, and market analyses for defense, intelligence, and space companies. His engagements assisted client's identification and quantification of new or evolving markets, described the political and business environments in which they operated, and provided execution guidance to ensure the implementation of successful ventures. Mr. Lambert also worked closely with a number of leading venture funds, merchant banks and private equity firms in support of their national security portfolio companies.

While with DFI, Mr. Lambert also led the company's work with first-tier defense firms, financial institutions, and private equity organizations in merger and acquisition market advisory services. Throughout his time with DFI, Mr. Lambert was engaged in scores of acquisitions, international joint ventures, offset arrangements and foreign equity investments, representing a wide variety of both buyers and sellers in the national security space.

Before joining DFI, Mr. Lambert worked for the US Department of State (USAID) at the American Embassy in New Delhi. Prior to this, he attended graduate school at Jawaharlal Nehru University on a Rotary Graduate Scholarship that he received during his senior year at Kansas State University. He also worked as an independent journalist in India, Pakistan and Burma. Before his time in Asia, Mr. Lambert served in the Political-Military Group at the Center for Strategic and International Studies.



Requisites for a Strong Defense Industrial Base

Lawrence P. Farrell, Jr.—Lt. Gen. Lawrence P. Farrell, Jr., USAF (Ret.), retired from the Air Force in October 1998 after 33 1/2 years of active duty. Prior to his retirement from the Air Force, General Farrell served as the Deputy Chief of Staff for Plans and Programs, Headquarters US Air Force, Washington, DC. He was responsible for strategic planning, resource programming and manpower activities within the corporate Air Force, and for integrating the Air Force's future plans and requirements to support national security objectives and military strategy.

Previous positions include Vice Commander, Air Force Materiel Command, Wright-Patterson Air Force Base, OH, and Deputy Director, Defense Logistics Agency, Arlington, VA. He also served as Deputy Chief of Staff for Plans and Programs at Headquarters US Air Forces in Europe. A command pilot with more than 3,000 flying hours, he flew 196 combat missions in Southeast Asia and commanded the 401st Tactical Fighter Wing, Torrejon Air Base, Spain. He was also the system program manager for the F-4 and F-16 weapons systems with the Air Force Logistics Command, Hill Air Force Base, UT.

General Farrell is a 1965 graduate of the Air Force Academy with a BS in engineering and an MBA from Auburn University. Other education includes the National War College and the Harvard Program for Executives in National Security.

A native of Montgomery, AL, he is married to the former Victoria Leigh Kruzel of Richmond, VA. They have a son, Lt. Colonel Sean Farrell, a daughter, Kelly Farrell Lowder, and 4 grandchildren.

Abstract

A vibrant industrial base is critical to US national security, for both economic and national defense reasons. Economically, the US industrial base is directly responsible for about 12% of US GDP, but it contributes as much as one third of GDP when considering commodities and services consumed. Further, over 60% of total US exports are manufactured goods, and 10% of total employment is within the industrial base. For reasons of national defense, the US must have full control of the supply chain that delivers supplies and equipment to those securing our nation.

Although the US has the largest industrial base in the world, there are concerns that must be addressed to maintain and strengthen our position. The often-stated but erroneous opinion that US manufacturing cannot compete globally, based upon labor rates, ignores the productivity gains from technology and the costs to move offshore. Other critical issues include the future availability of skilled workers, reliable supply of affordable raw materials, access to capital, uncertainty of the defense budget, and changing environmental policies. A key thread winding through all of these issues is the need for elected representatives to appreciate the vital role manufacturing provides. By addressing these concerns, the future of the US industrial base will be more vibrant, ensuring US security and a strong economy.



The Defense Budget and Defense Industry Finance

David Berteau—David J. Berteau is a Senior Adviser and Director of the CSIS Defense-Industrial Initiatives Group, covering defense management, programs, contracting, and acquisition. His group also assesses national security economics and the industrial base supporting defense. Mr. Berteau is an adjunct professor at Georgetown University, a member of the Defense Acquisition University Board of Visitors, a director of the Procurement Round Table, and a fellow of the National Academy of Public Administration. He also serves on the Secretary of the Army's Commission on Army Acquisition and Program Management in Expeditionary Operations.

Roy Levy—Roy Levy is a Consultant with the Defense-Industrial Initiative Group at the Center for Strategic and International Studies (CSIS), where he specializes in financial aspects of the US defense industrial base. Before joining CSIS, Mr. Levy was a Policy Analyst with a New York City-based economic research firm and was a Fellow at the Colin Powell Center for Policy Studies between 2007 and 2009. He is the author and co-author of several published articles on international security issues. Mr. Levy holds a BA in Political Economy from the City University of New York.

Matthew Zlatnik—Matthew Zlatnik is a fellow with the CSIS Defense-Industrial Initiatives Group, where he focuses on how technological, industrial, and budgetary issues affect defense policy. Mr. Zlatnik previously spent 10 years in investment banking, primarily working with corporate clients in the telecommunications industry. He holds an MA in international relations from the Johns Hopkins School of Advanced International Studies (SAIS), an MBA in finance from the Wharton School of the University of Pennsylvania, and a BA in economics from Carleton College.

Abstract

The defense sector's fundamentals in terms of operating margin and cash flow return on investment (CFROI) are stronger today than at any point in the past two decades due to better cash flow management, operating efficiencies, and record US defense spending. However, the economic and business environment for the defense sector moving forward is likely to be more difficult because of the Federal budget deficit and the government's non-defense spending requirements. Likely changes in spending priorities have the potential to change the industry significantly. Assessing the vulnerability of the defense sector to potential market changes—both as a whole and within the various segments it comprises—is of the outmost importance.

Introduction

To evaluate the vulnerability of the industry to potential changes in market conditions, it is important to understand the financial drivers of defense companies and to have insight into why companies enter or exit the industry and why investors choose to fund the sector. In this annotated brief, we examine the defense sector's profitability, both in a historical context and in comparison to commercial peers. We choose to evaluate the profitability of the industry because of its importance to a sector's attractiveness to outside investors. Low profitability relative to other sectors will hinder an industry's ability to finance its operations. In the context of the industrial and technological base supporting defense, to the extent profitability shrinks, the industry may shrink as well, as companies and capital gravitate toward more lucrative sectors.

Our analyses show that while the overall fundamentals of the defense sector in terms of operating margin and CFROI are stronger today than at any point in the past two decades, company performance may vary due to size (measured by revenue) and



specialization (from products to service, electronics to heavy metal). The main question to consider going forward is how can companies preserve margins and, by extension, maintain their attractiveness to capital markets in an era of flat or declining budgets?

The CSIS Defense Index

To examine and analyze the defense sector, CSIS will utilize existing tools and develop additional methods. To reflect changes and capture the diversity of companies within the sector, we created the CSIS Defense Index. Our index is composed of 23 public companies with revenue ranging from \$114 million to \$42 billion, representing not only hardware and equipment firms but also the professional services sector. We also include several foreign companies with a significant presence in the US defense market. Our financial analysis takes a bottom-up approach, aggregating data for each of the companies in the CSIS Defense Index to obtain weighted totals or averages. Financial data are obtained from commercial services.

The Defense Budget

Defense spending is at an historic high, surpassing in real terms both the Vietnam War era and the Reagan buildup of the 1980s (Figure 1). Top-line budget today is 85% higher than the top-line budget in 2001 (constant FY2010 US\$).

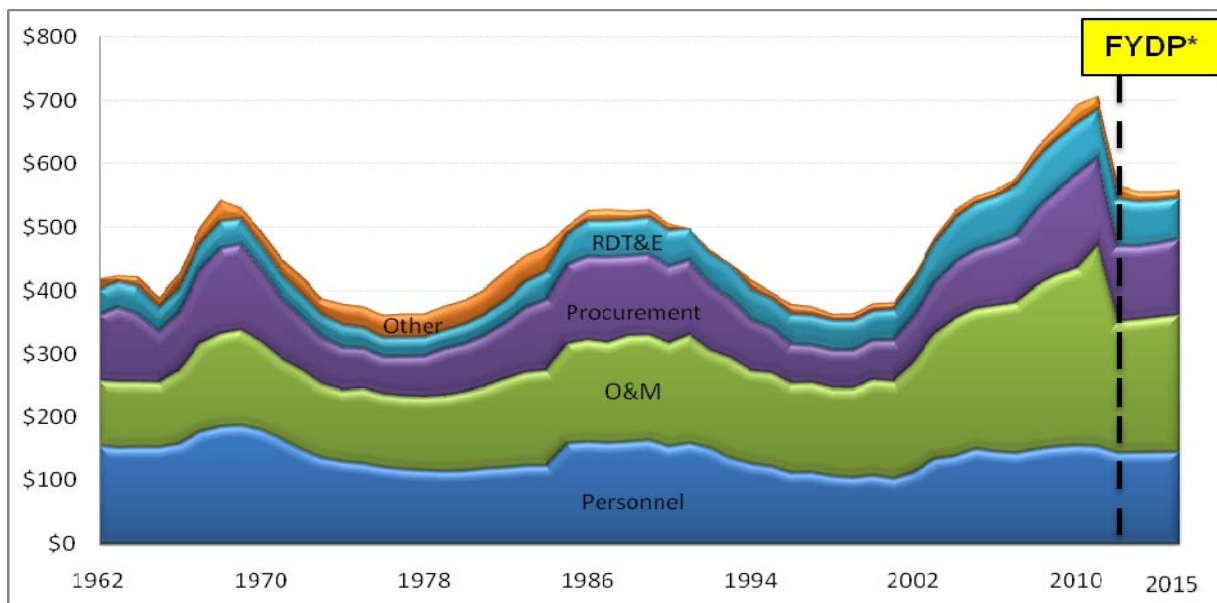


Figure 1. Department of Defense Outlays, 1962-2011 (constant FY2010 US\$ billion)**
(Source: DoD Comptroller; analysis by CSIS Defense-Industrial Initiatives Group)

*Excludes projected OCO

**FY1962-FY2009=Historical Outlays; FY2010=Budget Authority including OCO as of February 2011; FY2011=Budget Request Including OCO

From an industry perspective, while DoD contract action value more than doubled during the period, the relative composition of products and services procured remained largely intact (Figure 2). The one exception is the Professional, Administrative, and Management Support (PAMS), which grew from 26% of total professional services in 2000, to 34% in 2009.

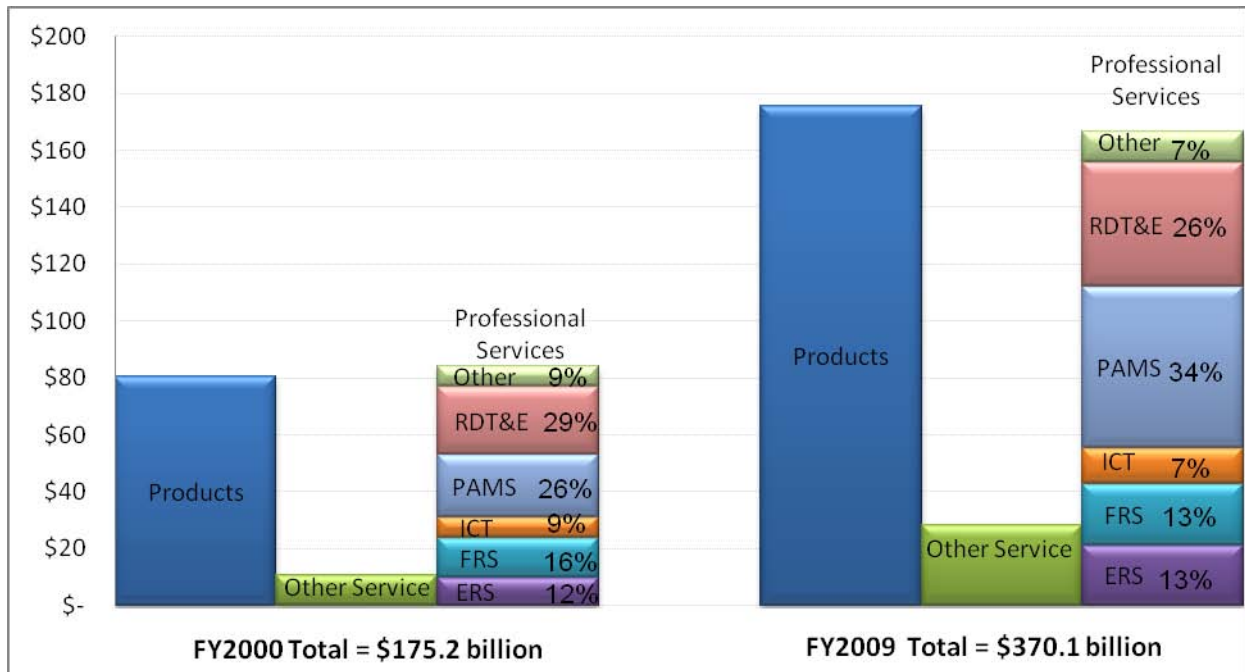


Figure 2. Composition of Total DOD Contracts, 2000 and 2009 (constant FY2009 US\$, billion)

(Source: Federal Procurement Data System, analysis by CSIS Defense Industrial-Initiatives Group)

PAMS: Professional, Administrative, & Support	FRS: Facilities Related Services
ICT: Information, Communications, & Technology	ERS: Equipment Related Services

Defense Industry Financial Performance

What does this all mean for the defense industry? From 2001 until the over-all stock market peak in 2007, the defense index outperformed the S&P 500. This outperformance was especially pronounced from March 2003 until late 2007, the period that saw the most intense fighting in Iraq. From 2007 to 2008, the indices traded closely together. Investors' anticipation of the end of US involvement in Iraq and slower-growing investment budgets could be one explanation.

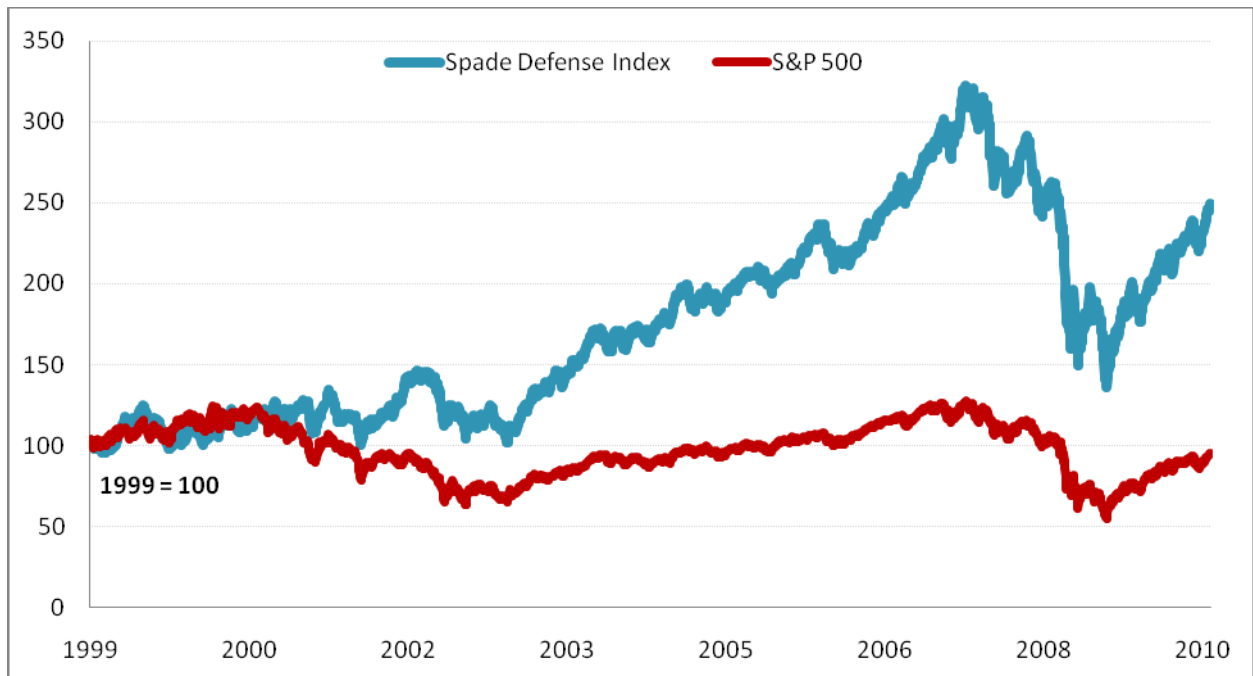


Figure 3. Defense Index Performance, 1999-2010 (1999=100)

(Source: Bloomberg, analysis by CSIS Defense-Industrial Initiatives Group)

The plunge in the Spade Defense Index in 2008 probably reflects broader investor flight from equity more than a fundamental change in the financial health of the defense sector. Since 2009, the defense index has outperformed the S&P 500, possibly reflecting greater investors' insight into the Obama administration's defense priorities, as outlined in the recently-released *Quadrennial Defense Review (QDR)*. It also is a testament to the relative health of the defense industry.

In the medium and long terms, profitability is critical to a sector's attractiveness to outside investors. Low profitability relative to other sectors will hinder an industry's ability to finance its operations. In the context of the industrial and technological base supporting defense, to the extent profitability shrinks, the industry may shrink as well, as companies and capital gravitate toward more lucrative sectors.

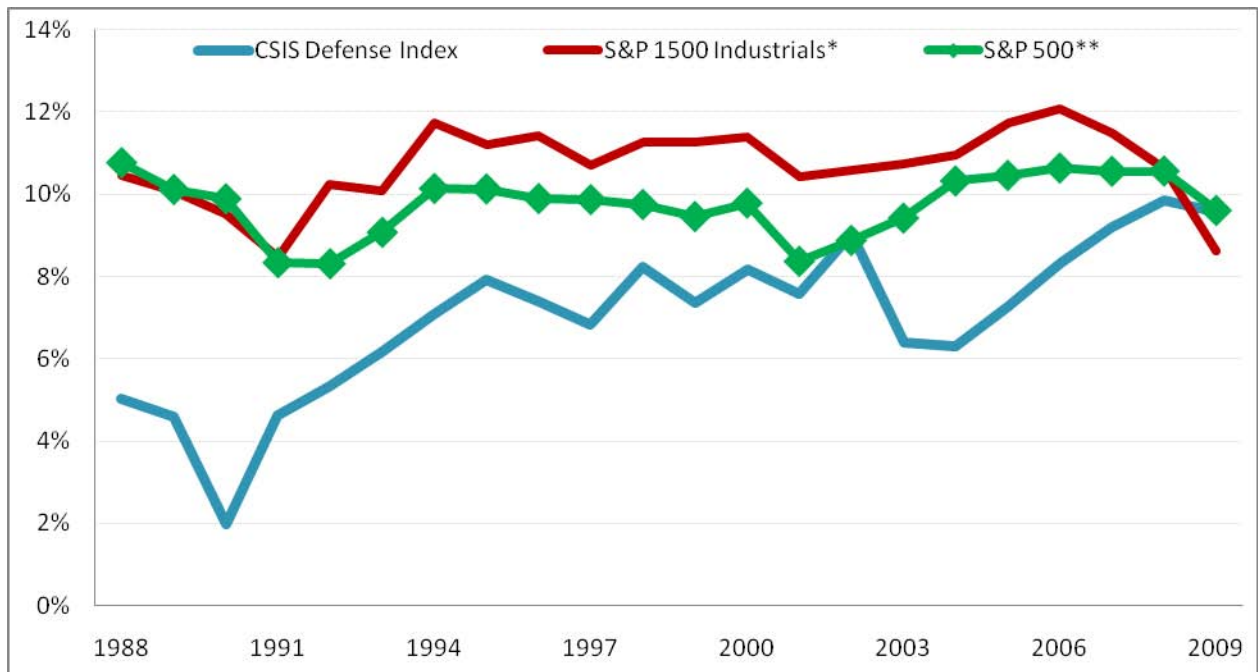


Figure 4. CSIS Defense Index Operating Margin, 1988-2009 (revenue weighted)
(Source: Bloomberg, analysis by CSIS Defense-Industrial Initiatives Group)

Earnings Before Interest and Taxes (EBIT) is a widely used measure of operating profitability. EBIT margin for the defense industry was consistently lower than its commercial peers during the period evaluated. 2009 marked the first year in more than twenty years in which operating margins for the defense sector were higher than that of the S&P 500 index, though this is due to a drop in the S&P 500's operating margin. Still, at nearly 10% for 2008 and 2009, the defense sector's total EBIT margin is at a high since 1988.

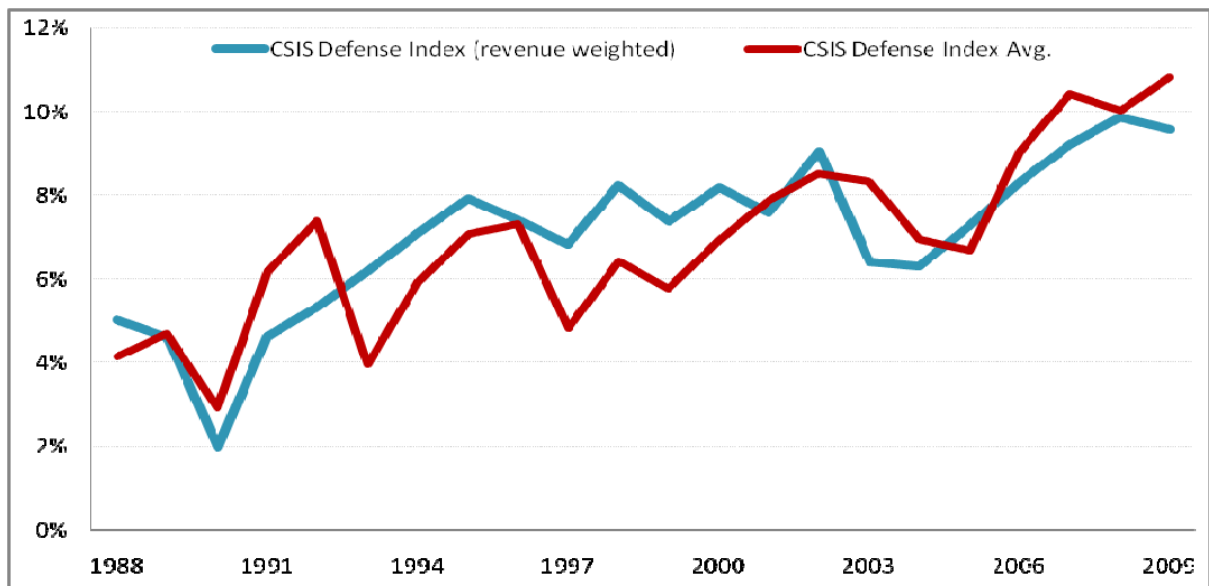


Figure 5. CSIS Defense Index Operating Margin, 1988-2009 Revenue weighted versus Index Average
(Source: Bloomberg, analysis by CSIS Defense-Industrial Initiatives Group)

However, as was mentioned in the introduction, the defense sector is not monolithic, and operating margin performance varies by company size (by revenue) and specialization. Figure 5 above shows that from 1993 to 2000, small and mid-tier companies in the CSIS Defense Index had lower operating margins than larger defense contractors.

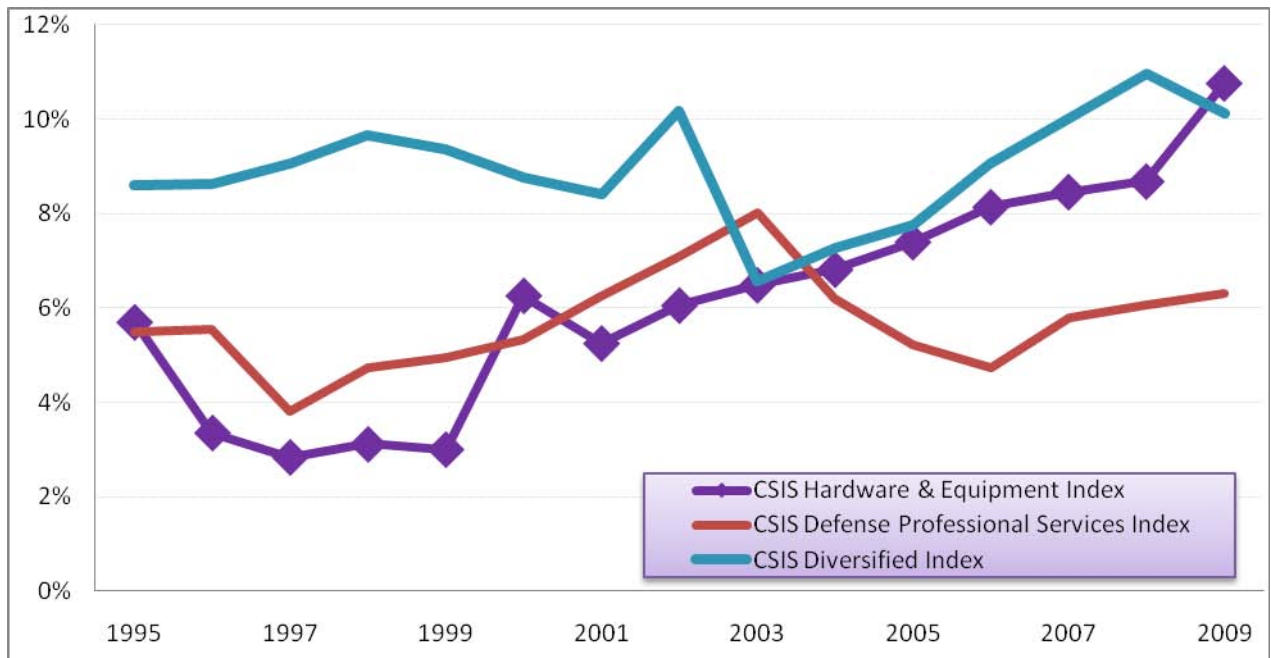


Figure 6. CSIS Defense Indices Operating Margin, 1995-2009 (revenue weighted)
(Source: Bloomberg, analysis by CSIS Defense-Industrial Initiatives Group)

When companies in the CSIS Defense Index are grouped by operating segments (professional services, hardware and equipment, and diversified), there seems to be an advantage for large or diversified companies. The drop in operating margin for professional services companies in 2003 is likely the result of increased competition in the segment.

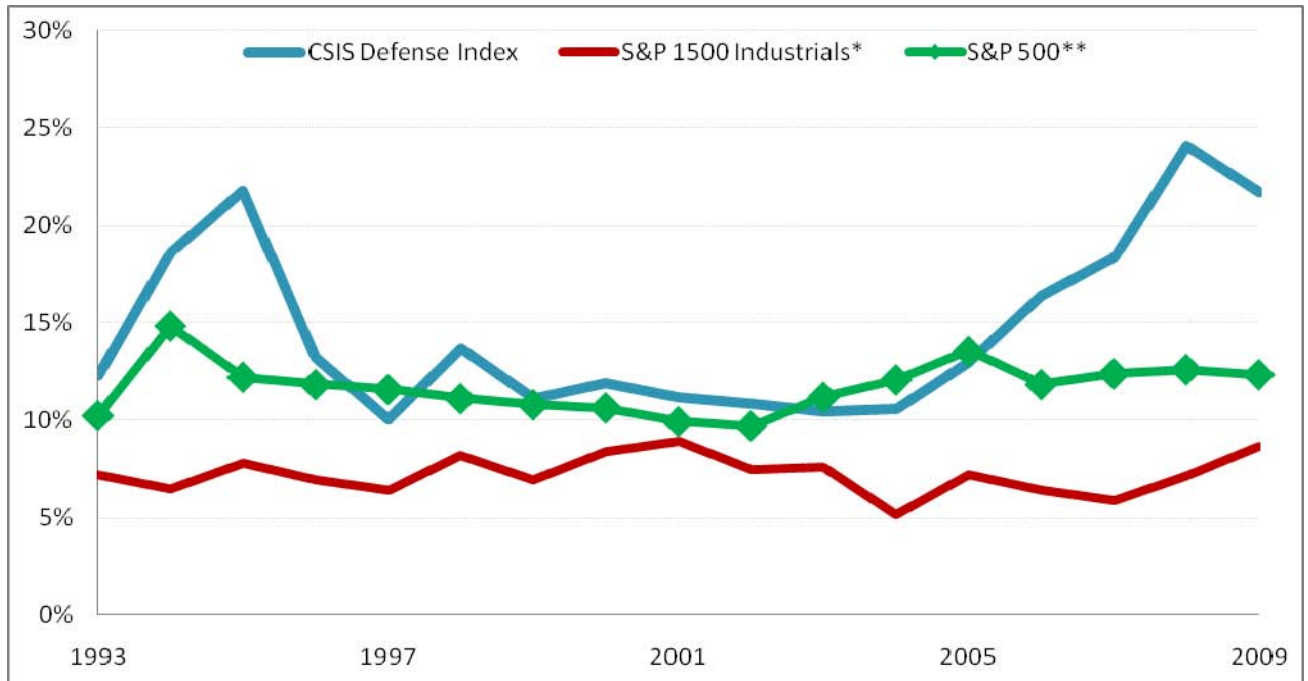


Figure 7. CSIS Defense Index Cash Flow Return on Investment (CFROI), 1993-2009 (revenue weighted)

(Source: Bloomberg, analysis by CSIS Defense-Industrial Initiatives Group)

A second, widely used profitability metric is Cash Flow Return on Investment (CFROI), which is a measurement of the cash flow available after expenses have been paid and sufficient investment has been made to continue current operations. CFROI is an important profitability measure to understand. Return on investment often drives decisions to enter or exit an industry, meaning it can ultimately shape the breadth and depth of the defense industry and the capabilities it offers. CFROI for the CSIS Defense Index has been higher than that for the indices of both the S&P 500 and S&P 1500 Industrial. The steep increase in CFROI, starting in 2004, can be partially explained by strong operating-income growth combined with stock buybacks that reduced total investment.

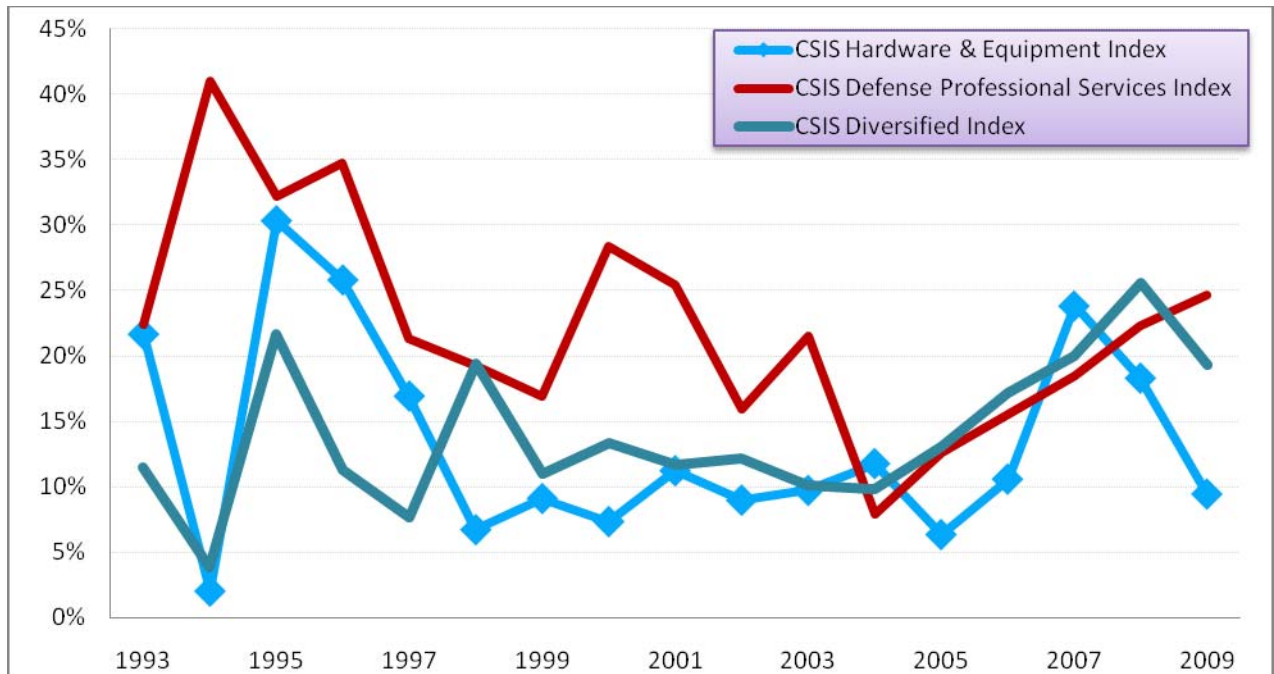


Figure 8. CSIS Defense Indices CFROI, 1993-2009 (revenue weighted)

(Source: Bloomberg, analysis by CSIS Defense-Industrial Initiatives Group)

Among the different segments, professional services companies have averaged higher CFROI than hardware & equipment and diversified companies, mainly due to generally lower net working capital. Beginning in 2003, CFROI for all three segments has tracked relatively closely together.

Summary

The overall defense sector's financial fundamentals are as stronger today than as they have been at any point in the past two decades. While operating margin has been lower for the defense sector than that of its commercial peers for the period, by 2009 operating margin for defense and commercial companies equalized. Within the defense sector, there appears to be an advantage for large or diversified companies. The defense industry had higher cash flows compared to its commercial peers for the period. Among defense companies, while professional services enjoyed relatively higher CFROI than hardware and equipment, or diversified companies, by 2005 CFROI for the different segments tracked closely together.

At this point, there are some key questions for industry and DoD policy-makers to consider. How can companies preserve margins (and by extension their attractiveness to capital markets) when their main customer is fiscally constrained? How can companies position themselves for an era of flat or declining investment? Where can companies find growth? What should the government rules and policies be with regard to industrial financial health?

THIS PAGE INTENTIONALLY LEFT BLANK



Panel #18 – Globalization: Cooperation and Competition in the Defense Industry

Thursday, May 13, 2010

11:15 a.m. –
12:45 p.m.

Chair: Dr. David Moore, Director, Centre for Defence Acquisition, Cranfield University

Discussant: Dr. Kevin Burgess, Senior Research Fellow, Centre for Defence Acquisition, Cranfield University

US Export Controls and Technology Transfer Requirements—A UK Perspective

David Moore, Pete Ito, Stuart Young and Peter Antill, Cranfield University

Global Cooperation and Competition in the Defense and Aerospace Industries

Raymond Franck and Ira Lewis, Naval Postgraduate School and Bernard Udis, University of Colorado at Boulder



US Export Controls and Technology Transfer Requirements—A UK Perspective

David Moore—David Moore worked in purchasing, logistics and supply chain management within public sector and commercial organisations before entering academia. He has designed, developed and delivered a range of professional courses, undergraduate and masters programmes for organisations and universities. He has undertaken extensive education, training, speaking and consultancy assignments in the UK, USA, Europe, Middle East and Far East. Particular interests include outsourcing, using contractors for service provision, and developing professionalism and humanitarian logistics. He has written a number of books, book chapters, and conference and journal papers. Moore completed his service in the Royal Logistic Corps as a Lt Colonel in 1999.

Dr. David M. Moore
Director, Centre for Defence Acquisition
Cranfield University
d.m.moore@cranfield.ac.uk
Office Phone: (44)1793-785-659
Office Fax: (44) 1793-785-871
Home Phone: (44) 1291-620-633

Stuart Young—Stuart Young retired from the Royal Navy as an engineer officer in 2008, having served in a variety of postings at sea and in the UK Ministry of Defence. These included three years based in the British Embassy in Washington as a technical liaison officer, and as programme manager for a major multi-national technology development programme. In his final appointment, he was responsible for the development of acquisition management skills for military and civilian personnel across the MoD. Joining Cranfield University in 2008, he is Deputy Director of the Centre for Defence Acquisition, with a particular interest in the relationship between the MoD and industry and the development of acquisition strategies for major defence programmes.

Stuart Young
Lecturer, Centre for Defence Acquisition
Cranfield University
s.young@cranfield.ac.uk
Office Phone: (44)1793-785-051
Office Fax: (44) 1793-785-871
Home Phone: (44) 1793-828-775

Pete Ito—Pete Ito earned a Bachelor's degree in Political Science from the University of California at Berkeley, and a Juris Doctor (law) degree and a Master's degree in International Affairs from George Washington University in Washington, DC. He worked for 25 years as a Foreign Service Officer for the US State Department, serving in South Korea, Denmark, Germany, The Netherlands and Washington, DC. His primary focus was political affairs, particularly defence and security policy. He joined Cranfield University in September 2007, working as a researcher in the areas of strategic management and change management, before moving to his current position.

Pete Ito
Lecturer, Centre for Defence Acquisition
Cranfield University
p.ito@cranfield.ac.uk
Office Phone: (44)1793-785-843
Office Fax: (44) 1793-785-871
Home Phone: (44) 1793-766-435



Kevin Burgess—Kevin Burgess has both private- and public-sector experience. In the past twenty years, he has held a range of senior management and executive roles in asset intensive industries, namely Telcos and Railways. Prior to joining Cranfield, Burgess's last job in industry in 2008 was as Group General Manager, Shared Services in QR (a railway with over A\$3 billion annual revenue and 15,000 staff). In this role, he reported directly to the CEO, had 985 staff and A\$600 million annual operating budget. His corporate-wide responsibilities included information technology, project management (capital program of A\$10 billion), learning and development, supply (A\$1 billion p.a.), HR services, financial services, property (A\$600 million in assets), and rollingstock engineering. He has ten years experience as a Business Excellence Evaluator (US equivalent of the Baldrige Award). His academic interests have been primarily on integrating social and technical systems in order to improve overall corporate performance. His PhD was in innovation in supply chains. He is widely published and currently holds the title of Adjunct Professor in three Australian Universities.

Dr. Kevin Burgess
Senior Research Fellow, Centre for Defence Acquisition
Cranfield University
k.burgess@cranfield.ac.uk
Office Phone: (44)1793-314-756
Office Fax: (44) 1793-785-871
Home Phone: (44) 1793-734-463

Peter Antill—Peter Antill is currently a research assistant working for Cranfield University at the UK Defence Academy's College of Management and Technology in Shrivenham. Peter has practical experience in the service industry as well as the civil service. A degree holder from Staffordshire University and the University College of Wales, Aberystwyth, he also holds a PGCE (Post Compulsory Education) from Oxford Brooks University. A published author, he is currently conducting post-graduate research into British defence policy, historical procurement programmes and expeditionary operations as well as privately collaborating with two colleagues in a military history project.

Peter Antill
Research Assistant, Centre for Defence Acquisition
Cranfield University
p.antill@cranfield.ac.uk
Office Phone: (44)1793-314-755
Office Fax: (44) 1793-785-871
Home Phone: (44) 1793-320-998

Abstract

It's not true that life is one damn thing after another; it is one damn thing over and over. - Edna St. Vincent Millay (1892-1950)

The paper and discussion address the international impact of US export control and technology transfer regulations. They highlight the current importance of this issue for the United Kingdom and outline an NPS-Cranfield University research project in this area. The focus will be on the experience of the UK with these US requirements in connection with the Joint Strike Fighter (JSF) program. This is relevant as the dispute regarding JSF source code access for the UK, resolved in 2006 at the President/Prime Minister level, was re-opened by the US in late 2009.



This is a key topic due to the White House announcement in August 2009 of an initiative to change the US export control regime. And it is a timely issue for the UK, as it prepares to conduct a Strategic Defence Review, including a review of international acquisition relationships.

The UK perception of its relationship with the US and the benefits of participation in US-led multinational projects will be influenced by its JSF experience. And the UK provides a unique perspective on the impact of US export control and technology transfer regulations, an issue of significance to those interested in multinational military cooperation with the US.

Introduction

At the 2009 Acquisition Research Symposium, the researchers from Cranfield University presented a paper titled Innovative UK Approaches to Acquisition Management. It addressed four areas in which the experience of the United Kingdom had relevance for practitioners of defence acquisition. One of those areas involved UK participation in the Joint Strike Fighter (JSF) programme. The difficulties that arose over the US refusal to grant the UK access to JSF source codes was an indication of the extent to which US export control and technology transfer policies generate obstacles to US-led multinational programmes. The point was made that other countries would probably ask themselves the following question: If this is the way the US treats what is generally accepted to be its closest ally, what should other countries expect? The 2009 paper could at least note that this was a historical issue, resolved in 2006, once it had been elevated to the Presidential/Prime Ministerial level.

However, it was too soon to declare resolution of this issue. As noted in a December 2009 report to Congress (Gertler, 2009, pp. 14-15), the US will not be releasing the JSF source codes, and, quoting the head of JSF international affairs, "that includes everybody." The official added, "Nobody's happy with it completely, but everybody's satisfied and understands" (p. 14). Instead, a reprogramming facility will be set up to develop JSF software and distribute upgrades. Changes to the software will be integrated at that facility "and new operational flight programs will be disseminated out to everybody who's flying the jet" (p. 15). The UK MoD statement from December 2009 in response to that announcement simply noted that the JSF "is progressing well and the UK currently has the JSF data needed at this stage of the programme, and is confident that in future we will continue to receive the data needed to ensure that our requirements for operational sovereignty will be met" (p. 15).

The resurrection of the source code issue does not bode well for supporters of multinational defence programmes, particularly those led by the US. And it comes at a critical time for the UK. With a general election likely to be held before the 2010 symposium, both the major political parties have promised they will hold a Strategic Defence Review (SDR) following the election, whatever the outcome. The last SDR was done in 1998. The MoD has been working to generate the papers needed to prepare for the SDR, and most, if not all, address the importance for the UK of a strong defence relationship with the US. As no project epitomizes the UK commitment to work with the US more than the JSF, the US decision to re-open the source code issue could not have come at a more sensitive time in the UK.

There is an appreciation in the UK that the US has every justification for protecting its cutting-edge military technology. However, what may not be fully appreciated in the US is the extent to which US export control and technology transfer policies have generated disincentives among allies and partners to participate in a US-led development programme.



And it is important to note that in efforts like the JSF, there is much that partner states can bring to the programme. For example, the UK is not just the only Level 1 participant in JSF, but BAE Systems is the largest sub-contractor.

The UK can provide a unique perspective on the impact of US policies, particularly due to the resurrection of the JSF source code issue. For that reason, the Naval Postgraduate School and Cranfield University will undertake a joint research effort over the next year into the impact of the US export control and technology transfer policies. The goal will be to provide the results at the 2011 symposium. The UK experience with the JSF can provide insights into the benefits and drawbacks of the US policies, which should be valuable to the US in pursuing multinational defence programmes.

Current UK Review of Security and Defence Policy

Green Paper. It is worthwhile to consider the extensive work the UK is undergoing to review its security and defence policy. This is particularly valuable due to the views that have emerged regarding the importance to the UK of its defence relationship with the US. A discussion of the impact of US export control and technology transfer policies should be placed in the context of this major UK review and the key documents that have been produced.

The recent MoD paper *Adaptability and Partnership: Issues for the Strategic Defence Review* (2010a), better known as the "Green Paper," notes that international partnerships and bilateral relationships will be important, "especially with the US" (p. 12). It goes on to comment that,

the UK has a range of close bilateral security and defence relationships. None is more important than that with the United States. The relationship is based on common values and interests which will endure in the 21st century, to our mutual benefit. The UK benefits greatly from bilateral co-operation in the nuclear, intelligence, science, technology and equipment fields. (p. 15)

The paper notably also comments that, "In Europe, the return of France to NATO's integrated military structures offers an opportunity for even greater cooperation with a key partner across a range of defence activity" (p. 15). This generated discussion in the UK, which has since abated. As indicated in the 2009 Cranfield paper, there is scepticism in the UK about the value of European programmes, such as the A-400M transport. However, it remains to be seen whether greater cooperation with Europe in general and France in particular will have any momentum when the SDR is undertaken. Certainly, to the extent that US decisions on the JSF or other programmes generates the view that the US is a difficult partner, this will buttress the case of those who support a greater emphasis on European cooperation.

One intriguing question for the UK is how the US would view a strengthened UK turn towards Europe. One of the issues the pending research will address is the perception of UK officials about the value the US places on UK contributions to the JSF. If the UK believes the US does not put great store in UK participation, reflected in the way in which it is treated, then such a relationship could have an impact on UK views of European cooperation.

The issue of acquisition receives considerable attention in the paper, and the commentary is worth reciting in full:

In 2005, the Defence Industrial Strategy (DIS) set out a comprehensive statement of how the Government would engage with industry on the acquisition of equipment, support, and services. The *DIS* will be updated during the future Review, in the light of future military capability requirements. We will have to revalidate our overall approach to:

- Operational Sovereignty. Our Armed Forces rely on assured overseas sources for some important equipment and support but there are cases where specific industrial capability must be located in the UK for operational reasons.
-
- International Collaboration. There are operational, industrial and economic benefits from working with other countries on acquisition. However such acquisition involves risks, constraints, and potential costs. We must choose the right approach for each project. But the Review must set guidelines.
-
- The broader benefits to the UK from our acquisition. The Ministry of Defence must provide the Armed Forces with the equipment they require at best value to the taxpayer. However, our annual global expenditure with industry and commerce—some £20 billion per annum—means that our decisions have a significant and long-term impact on the UK's industrial base and therefore on the livelihood of many of our citizens. (p. 38)

The paper addresses the issue of defence exports, a key part of deciding whether and to what extent the UK will commit itself to a defence development programme. The paper notes that,

Support for defence exports remains an important aspect of defence policy, in that it can reduce equipment costs to the UK tax payer, support jobs, facilitate bilateral defence links with allies and friends and enables countries to take responsibility for their defence and security needs. It is therefore in the MoD's interest to work with industry to take account of possible future exports when developing equipment for the UK Armed Forces. Longer term certainty on our future equipment requirements will also help industry plan their investment in new technology. (p. 38)

Finally, the paper addresses military technology. It is fair to state that the UK has been inclined to see the benefits of cutting-edge technology. However, the paper has a more nuanced view of the issue, influenced by the UK experience in Iraq and Afghanistan. Such a shift in perspective, particularly in an era of tight defence budgets, can have an impact on the willingness of the UK to commit scarce resources to expensive, high-tech military programmes:

Loss of our technological edge in significant areas of military capability would have a profound effect on the way we operate. For at least the last twenty years, we have operated on the underlying assumption that our equipment would be more effective than our adversaries'. If it were not, our operations would be more hazardous. Our casualty rates, in particular, could be expected to increase markedly.

We must also ensure we can bring technology to bear on the challenges we face. The most immediate threats may not be posed by the most advanced technology. The unique tactical threat posed by, for example an improvised explosive device or suicide bomber, can rapidly negate an assumed technological edge. We must be able to adjust our programmes rapidly to access the right technology in response.

A key challenge for Defence will be to monitor and respond to the increasing breadth and pace of technological change. We will need to develop a greater understanding of the requirement for technological edge in our systems and of the risks associated with losing it. We will need to be more agile in exploiting new technologies in our own capabilities. We need to recognise that the technology we require depends on the threat we face. (p. 23)

In short, the paper provides insights into some of the questions that will be addressed in the SDR, many of which will focus on the policy concerns regarding the UK-US security relationship. And the practical UK experience of operating alongside the US in Afghanistan, or in programmes such as the JSF, will have an impact on the policies that emerge from the SDR.

Acquisition Strategy. The MoD released *The Defence Strategy for Acquisition Reform* (2010b) to complement the Green Paper. The document notes the MoD spends £20 billion annually on goods and services, around two-thirds of the total Defence Budget (p. 6), with £6 billion on equipment and £5 billion supporting equipment in service (p. 7). The MoD reports that "nearly 90% of our equipment projects now deliver front-line needs to cost, and over 80% deliver them to time" (p. 6). However, the report highlights the challenges faced by the MoD, particularly in view of tighter budgets:

The difficulties we face in this area were considered in the independent report into Defence acquisition by Mr. Bernard Gray, which we published in October 2009. His sobering analysis, subsequently echoed by the National Audit Office (NAO), was that our overall plans for new equipment are too ambitious, and need to be scaled down to match the funding likely to be available. He also concluded that we must continue to improve the way we manage our equipment projects, both individually and as a portfolio. We accept this analysis and, building on previous reforms, have framed much of this strategy around it. (p. 7)

On the specific issue of international cooperation, the strategy outlines the key factors that should be considered, noting that

Deciding whether to acquire equipment in collaboration with other countries has crucial ramifications for its performance, cost and timescale. It often has important implications for international Defence relationships more generally, and for the Government's wider foreign and security policies. We need to make sure all these issues are properly weighed in reaching final acquisition decisions. (p. 20)

With regard to technology, the MoD states that it will be important to "examine the scope for managing technology and innovation better so that we can provide and update defence equipment more quickly, and at a price we can afford. This includes using incremental and modular approaches so that we can keep pace better with the evolving threats" (p. 10). The strategy also sets out a measurable standard for success in acquisition:

But the overall success—or otherwise—of the strategy will be visible in more transparent ways. Two clear indicators will be that, following the planned SDR, we can:

- demonstrate each year, with independent audit, that we can afford our equipment and support plans (Chapter 2); and
- demonstrate much lower levels of cost growth and delay across our equipment programme. We have set hard targets that, on average, cost growth should not

exceed 0.4% a year, and slippage should be under 0.8 months. We have also set targets to ensure the large majority of projects are managed within tolerable cost and time limits each year. (p. 8)

While a program like the JSF may not be bound by the MoD's acquisition strategy, the document, like the Green Paper, provides a good insight into the key issues being considered by the MoD. As is probably the case with most military forces, the MoD's ability to tolerate cost overruns and programmatic disputes on issues like technology transfer has largely evaporated.

Strategic Trends. The third key MoD document is its *Global Strategic Trends* paper (2010, January 12). On the general topic of US export control and technology transfer policy and the specific issue of the JSF, the paper has particular significance due to its assessment of future trends. It predicts that "the hegemonic dominance of the U.S. will fade," and that "she is likely to remain the pre-eminent military power, although, in political, economic and military terms, she is likely to be increasingly constrained as others grow in influence and confidence" (p. 10). The paper predicts that "By 2040, the U.S. is likely to lose her hegemonic status as rising powers enjoy more rapid economic growth and close the technology gap in military capability" (p. 45).

However, the paper clearly indicates that while the balance of military power will become multi-polar, the US is likely to remain pre-eminent (p. 80). And there will remain, at least out to 2020, a global expectation that the US will provide international leadership in times of crisis (p. 46). Most pertinent for this discussion, the paper opines that while most developed countries will minimise defence expenditures, the US is likely to be the exception, "making by far the greatest commitment to defence, although its economic power and technological advantage is likely to become increasingly challenged" (p. 80).

The report notes that "Defence production is likely to become increasingly internationalised and most states will lack guaranteed access to industrial surge capacity during times of escalating tension" (p. 17). This is an issue that will become increasingly important for the UK, particularly in light of the need for supply chain agility. With regard to the JSF, it will be one of the topics the Cranfield researchers will be addressing from a UK perspective. And the specific chapter on military research and development has particular relevance to the focus of this paper and warrants extensive citation:

R&D funding can be divided into 3 broad categories: private sector; government non-military; and military. The first of these is much larger than the others and is likely to grow. However, it is increasingly likely that defence and government budgets will be unable to service the totality of the defence and security R&D need; novel approaches to address the shortfall will be sought. The development of specific military technologies will out of necessity remain largely a government activity. For the remainder, the industrial base will be stimulated through 'seed corn' initiatives that promote development of novel technologies. Other avenues that are likely to be pursued include forming international, value-adding partnerships in military R&D. These are likely to sustain and acquire key enabling technological knowledge and capability, pull through technologies from multiple sources, particularly civil R&D, and harness the capabilities of academia and other civil research institutions. However, even where the civil sector is the dominant driving factor, transforming non-defence technologies into military advantage may require significant expenditure in defence R&D. (p. 136)

Returning to the issue of Europe and its future development in the security arena, the report judges that,

Within Europe, the EU is likely to increase its influence and expand its economic, foreign policy and security role. However, a more extensive defence relationship that would extend EU power into and beyond Europe's near abroad is unlikely. For European powers, NATO is likely to remain the defence organisation of choice. (p. 44)

And specifically with regard to further European defence integration, the assessment is that,

Further coordination of military forces is likely. In particular, financial restraints are likely to result in a requirement to pool and share capabilities, for example, capabilities such as strategic airlift. NATO is likely to remain the guarantor of European security, despite the fact that the U.S. will be less focused on Europe. The EU is likely to remain reluctant to project military power beyond the Petersberg tasks even in cases of clear multilateral interest such as in the Balkans, or where the humanitarian imperative is clear. (p. 48)

In short, the MoD prediction is that while the US will no longer have as much of a lead over all other nations as it has now, it will still be the pre-eminent military power. And while Europe may develop more capabilities, that is not synonymous with sufficient capabilities. Finally, the requirements for defence acquisition programmes will become more complicated, even without the added burden of an onerous export control and technology transfer regime.

Character of Conflict. The fourth MoD document concerns the *Future Character of Conflict* (2010, February 12). On defence acquisition and the importance of strong ties to the US, the paper addresses two points. The first is that in the key period around 2020, "the UK's strength in defence technology, along with many other Western nations (but probably excluding the U.S.), may have been surpassed by these emerging powers," emphasizing the likelihood is that UK defence budgets will be tight in this period (p. 29). The opinion in the paper is that "investments now in technology could hedge against relative technology decline in UK defence" (p. 29).

The second point is commentary of the centrality of the US for UK military planning. One of the key assumptions is that "the UK will act with others where shared interests and values coincide. We will routinely operate with allies and partners, in particular as a supporting partner in a U.S.-led coalition," adding that "it is extremely unlikely that the UK will conduct warfighting without U.S. leadership, but in other operations the UK may be called upon to lead a non-U.S. coalition" (p. 3).

Bernard Gray Report. The report from Bernard Gray (2009) was commissioned by the MoD, and intended to provide an independent assessment of UK acquisition policy. It has become an important part of the UK discussion on defence acquisition. While not coming to any conclusion regarding the benefits to the UK of participation in multinational projects like the JSF, Gray provides the following commentary:

A number of UK acquisition projects that are currently underway involve collaboration with other nations to jointly procure equipment for use by the military forces of all of the participants. The potential economic benefits of acquiring equipment in this way are significant:

- fixed development costs can be defrayed amongst a number of partners, generating economies of scale and associated reductions in unit costs that would not be otherwise be realised. This has led the UK to participate in some smaller

scale collaborative projects, including NLAW (light anti-armour missile), GMLRS (guided multiple launch rocket system); and

- projects that are very large and which could not realistically be undertaken independently by the UK; in such circumstances, collaboration affords access to technologies and capabilities that would otherwise be out of reach. This means the set of collaborative projects in which the MoD is involved are high-profile such as Typhoon, A400(M) and PAAMS on the Type 45 destroyer.

The potential benefits of collaboration are most evident on large, expensive projects with significant technical challenges to be overcome. Equally, these projects tend to have a high profile; any delay or overrun on these projects is likely to be very visible and embarrassing to the MoD.

The inherent difficulties in ensuring that all participants in any collaboration have their interests aligned is widely held to be at the root cause of many problems and, more generally, the view across the MoD and the wider defence industry is that such problems are a characteristic of all collaborative projects to a greater or lesser extent.

It is certainly the case that the divergence between the objectives of the various partners has led to problems on some projects. However, the question of whether the poor reputation of collaborative projects is warranted across the board remains open. This study has not examined the relative performance of collaborative projects in detail; the small sample and the specific issues raised in relation to each project render any such analysis relatively meaningless on a statistical basis. (p. 83)

Haddon-Cave Report. While not directly relevant to the issue of export controls and technology transfer, another document has relevance to the general discussion topic. The report from Charles Haddon-Cave (2009) focussed on the loss of a NIMROD aircraft in Afghanistan in 2006. The task for Haddon-Cave was to address the problems that had developed in MoD acquisition that led to a weakening of the attention to safety. However, one of the underlying themes of the report was the need for the MoD to be an intelligent customer, and not find itself at the mercy of industry and suppliers with whom the MoD were not able to properly engage.

A further point noted by Haddon-Cave was the need to control the supply chain. Again, this was under the overarching issue of the MoD's need to manage the acquisition process. But the issue has relevance regarding the extent to which the UK must be in a position to manage complex multinational programs, such as the JSF. One final theme is the need to manage capabilities through-life, and a focus on details after the initial procurement is completed. One of the issues with regard to the JSF is that the Lockheed support solutions for the F-35 may not be fully compatible with the systems the UK has in place. As the UK has learned from other acquisitions, as well as Iraq and Afghanistan, the task is not simply providing materiel to the central distribution point, but being able to transport it that last mile. This is all the more significant for cutting-edge fighter aircraft deployed at sea.

UK Participation in the Joint Strike Fighter

Background. Having placed the JSF and the specific issue of US export control and technology transfer policies in the context of the UK policy review, it is now time to turn to the JSF and UK involvement. A short review of the background of the JSF, the largest

multinational development program ever and the DoD's largest procurement program in terms of total acquisition cost, is worthwhile, as is the genesis of UK participation in the JSF.

The DoD plans for acquisition of over 2,400 JSFs for the Air Force, Marine Corps and Navy for \$246 billion, with hundreds of F-35s to be purchased by US allies (O'Rourke, 2009, p. 1). In the 1990s, the US Navy, Marine Corps and Air Force were working on a next-generation strike aircraft. In 1996, the JSF project was announced by the US. At about that time, the British Royal Navy was looking at new Future Carrier Borne Aircraft capability for its CVF programme. The US and UK decided to combine efforts. The requirements of the US services drove the program, but the UK focus was on an attack aircraft with advanced Short Takeoff and Vertical Landing (STOVL) capabilities so that it could operate from forward battlefields as well as from aircraft carriers. The UK preference for the JSF was confirmed in a 2001 MOU with the US. In 2002, the UK selected the STOVL variant to meet its future requirements, with a positive review of the JSF programme and the STOVL design completed in 2005 (National Audit Office, 2008, p. 45).

It was important that as an indication of US support for this endeavour, the Office of the Secretary of Defence sent down instructions that the JSF program should emphasise international participation, and there was a consensus that the UK would participate in the program (Franck, Lewis & Udis, 2009). The US and UK engaged in extensive, detailed exchanges on the nature of the program and the UK role, compiled in the US-UK Engineering and Manufacturing Development Framework Memorandum of Understanding. Comprised of agreements, letters and other supporting material, it provides the details of the US-UK relationship on JSF.

While decisions on specific numbers of fighters to be purchased were not required at the outset, the Royal Navy and Air Force were looking at the purchase of some 150 STOVL fighters to replace the Harriers. The UK participated from the outset of the program, and is the only "Level 1" partner contributing \$2 billion to the system design and development phase (Bolkcom, 2009, February 17, p. 9). That designation means the UK has significant access to most aspects of the programme as well as the ability to influence requirements and design solutions. And the UK will not have to pay the non-recurring R&D cost recoupment charge that normally accompanies the purchase of US military equipment and will receive a share of the levies on sales to third parties.

However, the UK commitment to the JSF is not primarily based on programmatic considerations such as cost savings, but on a key national security determination. The UK made a policy decision on the need to retain an aircraft carrier capability, and the requirement for carrier-based fighters derived from that critical decision. The postponement of the aircraft carrier production schedule by another two years due to constraints on the defence budget mean that the two carriers now have projected in-service dates of 2015/16 and 2016/17. However, that decision on the carriers is separate from the discussions on the JSF fighters that will operate on those carriers.

There is no doubt that the UK has made a significant wager in tying itself to the JSF program, for if there are major problems with the fighter, the UK will need to generate a "Plan B" to avoid having its aircraft carriers merely serving as floating platforms. If the STOVL version of the JSF does not emerge, then the UK would be faced with the challenge of re-designing ships so that, for example, they would have new catapults. The UK plans to bring in a total of 138 JSF fighters, with 60 to be acquired in the period 2015-19 (Joint Strike Fighter, 2007). The MoD is already in the process of working through modelling and simulation to optimise the safety and operability of the new aircraft carriers and the JSF when the fighters arrive (Scott, 2009).



The delays and cost increases in the Eurofighter program were a factor in the UK deliberations on whether to participate in JSF. In the aftermath of the Eurofighter experience, the fact that the JSF was structured so that the most competitive firms would win contracts was appealing to the UK, and in line with the goal of pursuing more efficient acquisition programs. And the fact that the US was providing the overwhelming amount of funding for a program with cutting-edge military technology was significant to the UK.

Adding to the list of factors pressing the UK to participate in JSF, the Defence Industrial Strategy emphasises the need for the UK to remain interoperable with Allies, particularly the US. And British industrial participation amplifies the UK focus on the JSF. BAE Systems is the largest non-American participant in JSF and has hoped for around £14 billion in development and production contracts (McGhie & Gee, 2006). Such a high level of BAE participation is to be expected, as it does the majority of its business in the US and is one of the largest suppliers to the DoD. Indeed, BAE participation in the JSF was viewed in the UK as a seal of approval on the British ability to participate in cutting-edge military projects. And there are potentially significant economic benefits. Bolkcom (p. 17) notes that the DoD conducted a 2003 assessment that determined that partner nations could potentially earn between \$5 and \$40 of revenue for every \$1 invested in JSF program contracts.

It is also important to note the extensive US-UK military cooperation as the backdrop for discussion of the JSF source code dispute. Franck, Lewis & Udis (2009) note that it is estimated that 99.8% of licenses for UK-US transactions are approved, which accounts for some 8,500 items with a value of \$14 billion (p. 88). This indicates that the routine operation of bilateral defence cooperation and technology transfer proceeds without friction. Moreover, the extent of UK-US defence industrial interconnection has increased substantially. Franck, Lewis, and Udis (p. 98), citing Chao and Niblett (2006), note that aside from BAE, UK firms have acquired 50 aerospace and defence firms in the US since 2001, which constitutes some three-quarters of all foreign investment in the US defence sector. And major American defence contractors are established in the UK or have acquired operations or set up a presence in the UK.

JSF Source Code Dispute. With this backdrop of strong bilateral defence cooperation, the specific problems that arose regarding UK access to JSF source codes generated doubts in the UK regarding US-led military cooperative efforts that should (and could) have been avoided. Initially, UK commentary on the JSF was full of praise as a model for future multinational defence cooperation. That turned to criticism of the JSF as an example of why such efforts may not pass an all-encompassing cost-benefit analysis.

It is a key operational requirements as well as a matter of sovereignty for the UK to be able to have the information needed to integrate, upgrade, operate and sustain the JSF as required. As a practical matter, the UK cannot buy into a system that requires a US maintenance team to take care of any problems that may arise or to arrange for required modifications. The House of Commons Defence Committee (2005) reported:

It is vital that the UK gets all the information and access to technology it requires from the U.S. to have 'Sovereign Capability'—the ability to maintain the Joint Strike Fighter aircraft and undertake future upgrades independently. The UK must receive adequate assurances that it will get all the information and access to technology it requires before the programme is too far advanced. If these assurances are not given, it is questionable whether the UK should continue its involvement in the programme. (p. 3)



The Committee (2005, p. 29) emphasised the UK could not accept a situation in which it could not operate the JSF independently of the US and pressed the Defence Minister to ensure the UK would have operational independence. It noted its expectation that the MoD would set a deadline by which the assurances on sovereign capability would be obtained from the US. In December 2006, as the source code issue was heating up, the Committee warned that an assurance from the US was needed by the end of the year that it would provide the UK with all requested technical information. In the absence of an agreement by the end of 2006, the Committee called on the government to develop a “plan B” to obtain alternative aircraft (BBC News, 2006).

From the UK perspective, the history of the political discussions to resolve the source code issue is not the best advertisement for multinational programs. The fact that such issues never seemed to be fully resolved added to UK frustration with US policy. The technology transfer dispute had been raised in 2004, when Secretary of State for Defence Hoon wrote to Secretary of Defence Rumsfeld and referred to the fact that the US had signed an outline agreement on defence technology cooperation in 2002 (O'Connell, 2004). Prime Minister Blair believed he had reached an agreement with President Bush in May 2006, but the dispute lingered on unresolved until the end of the year (Baldwin, 2006). Indeed, in the US *Defense Authorization Act for Fiscal Year 2007* (2006), Congress, aware of UK concerns over this issue, flatly wrote that “It is the sense of Congress that the Secretary of Defense should share technology with regard to the Joint Strike Fighter between the United States Government and the Government of the United Kingdom consistent with the national security interests of both nations” (Section 233).

The importance to the UK of technology transfer in cooperative arrangements with the US was already set out in the MoD's *Defence Industrial Strategy (DIS)*, in which the significance of the US defence market and US defence spending was acknowledged. The *DIS* commented:

To meet our own sovereign needs, it is important that we continue to have the autonomous capability to operate, support and where necessary adapt the equipment that we procure. Appropriate technology transfer is therefore of crucial importance. This is so for any cooperative project, but in practice difficulties have arisen particularly with the U.S., whose technology disclosure policy we have found less adapted to the needs of cooperative procurement than those of our partners in Europe. To reiterate, this is not about gaining competitive advantage for UK industry; it is about being confident that the equipment we buy meets the capability requirements against which it is procured and can be modified effectively to meet emerging requirements through life. We fully recognise the need to ensure that intellectual property is protected, and that appropriate measures are put in place to ensure this; security is a key issue for us, just as it is for the USA. But a certain degree of technology transfer is required if we are to be able to fully cooperate with the USA (or any other partner) on our equipment programmes. What we are striving towards is an agreed framework which facilitates this whilst ensuring that our mutual security needs are met. (p. 45)

Frustrations with US policy are exacerbated when problems arise even when sensitive technology is not involved. One Lockheed Martin employee noted that the restrictions on technology transfer have been “far more cumbersome and impenetrable than originally envisioned” and it is necessary to ask for Washington’s approval of “even unclassified information exchanges” (Metz, 2005, p. KN3-7). Other close allies such as Australia have found the American approach extremely trying, with Australia, at about the

same time, pushing "for a resolution of long-standing technology release issues with the USA" (La Franchi, 2006). Such requirements make it difficult for partners to participate and generate a large administrative burden on team members, who face the requirement that "all information is releasable under penalty of jail terms—not a conducive atmosphere for co-engineering a product" (Metz, 2005, p. KN3-7). And this has an impact on the UK evaluation of whether participation in US-led programmes is possible, and if the benefits outweigh the costs. In late 2006, with the issue of source codes at its most contentious, an unnamed UK MoD official was quoted as stating, "If we can't trust the Americans to provide this, then you would have to ask what else we should be doing with them in defence terms" (Baldwin, 2006).

It is worth noting the report of the Inspector General (IG) of the DoD (2008) on security controls regarding JSF classified technology that assessed seven applications from Northrup Grumman and BAE Systems for detailed review. The IG's office also evaluated security reports on BAE facilities. While the assessment from the IG's office was a frank statement that JSF advanced technology "may have been compromised by unauthorized access at facilities and in computers at BAE Systems" (p. ii), the specific criticism was that the DoD did not always use sufficient controls to evaluate potential unauthorised access to such technology (p. i). Indeed, the specific recommendation with regard to BAE Systems is that the Defence Security Service (DSS) could have bolstered its efforts by collecting, analysing and retaining security audit reports completed by BAE Systems, a point on which the Director of DSS concurred. And the other recommendations involved the actions of DSS.

JSF Source Codes—The Sequel. After the source code issue was resolved in 2006, the JSF became a low-profile project. The House of Commons Defence Committee (2009) simply noted that the MoD has assessed that the JSF programme is "progressing well" and the Committee would monitor the progress of the program (p. 47). Secretary of State for Defence Hutton announced in March 2009 that the UK would purchase three F-35B operational test aircraft, indicating the UK commitment to the Operational Test and Evaluation phase of the JSF (JSF, 2009).

However, the announcement in late 2009 from the JSF project office that source codes will not be released has re-opened the issue. And while the MoD response avoided a direct challenge on the matter, there is every reason to anticipate that UK objections raised the first time around will be reiterated. The House of Commons Defence Committee commented in its report *Defence Equipment 2010* (2010), "We also note that there still appear to be outstanding issues concerning technology transfer for the JSF, which are of key importance to the success of the programme" (p. 25).

The second edition of the source code debate may be more heated than the first, as key UK concerns have been thrown into the mix. The US discussion on the JSF is centered on the Administration's push to cancel the second engine (the F136). As the F136 is a General Electric/Rolls Royce effort, the UK has an economic interest in the future of the second engine option. As Bolkcom (2009, February 18) noted to Congress, the UK's senior defence procurement official had stated the UK would discontinue participation in the JSF if the technology transfer issues were not resolved and if the F136 engine were cancelled (pp. 9-10). Dr. Liam Fox, currently the Conservative party shadow defence secretary, said in 2006 that a US decision to drop the F136 would "invariably effect future procurement decisions, with seriously negative consequences that may not be fully appreciated on this side of the Atlantic" (Cahlink, 2006; Warwick, 2006). In short, while not an export control or



technology transfer issue, a debate on JSF source codes will probably be intensified if the second engine is deleted.

An added factor is the increase in the cost of the JSF. The UK defence budget is expected to come in for some cuts, irrespective of the outcome of the general election. The cost increases that have been noted for each JSF (from £37 million four years ago to approximately £62 million today) have generated speculation about whether only one aircraft carrier might be fully fitted out with F-35s. Other options include upgrading older aircraft, buying "off-the-shelf" competitors (such as the Rafale M or F/A-18E Super Hornet), using the CV version of the JSF instead of the STOVL version, at least on the carriers, or altering one of the carriers to an "assault" role, equipped with helicopters, Royal Marine Commandos and UAVs (Norton-Taylor, 2010). While all of this is still in the realm of speculation, the fact that such options are being considered are critical, as they arise from the fact that there is no prospect of savings with regard to the two carriers. The only savings can come from the aircraft that will be placed on those carriers.

A final consideration is the fact that the US-UK Defence Trade Cooperation Treaty, signed in 2007, is still on hold. The House of Commons Foreign Affairs Committee (2010, p. 3) noted that "We are disappointed that despite promises to do so, the U.S. Senate has not yet ratified the UK-U.S. Defence Trade Cooperation Treaty. We conclude that its swift ratification is imperative and would bring a range of benefits to both countries." The goal of the treaty, to cut red tape in the bilateral exchange of defence goods, services and information, would be a productive step forward. It is an open question whether the treaty, had it been in place, would have had any impact on the resurrection of the JSF source code dispute. What is clear is that the delay in ratification indicates to the UK that improving the bilateral defence relationship is not a top priority for the US. And it is notable that the Committee, although not referring specifically to the JSF, stresses that the UK should continue to work closely with the US but that "the UK needs to be less deferential and more willing to say no to the U.S. on those issues where the two countries interests and values diverge" (p. 7).

It is worth reiterating that the resumption of this US-UK dispute has an impact on the perceptions of other nations regarding the attractiveness of US-led military development programs. As noted by Chao and Niblett (2006):

if the United States and the UK, the two closest of allies, are unable to overcome the continuing obstacles to the efficient sharing of defense-related technologies, what hope is there for broader transatlantic defense industrial and technological cooperation? Bilateral U.S.-UK cooperation in the fields of intelligence, nuclear defense, and military deployments is unprecedented in U.S. alliances. And the U.S. and UK defense industrial bases have become increasingly intertwined through investment and trade. And yet, the U.S. and UK governments have proved unable to institute a more open system for exchanging and transferring defense technologies, despite the stated intent of senior political leaders and extensive efforts by both sides over the past couple of years. (p. 3)

The authors stress that there is substantial willingness on the part of the UK to make this bilateral defense relationship work. The problem is the lack of a corresponding effort on the part of the US. As the authors note,

The bottom line is that UK political leaders remain committed to working with the United States as the closest of security allies. They want to be capable to deploy and operate UK forces from the outset, shoulder to shoulder with U.S. forces in the most intense and complex of battlefield situations. They also want to take advantage



of U.S. defense equipment and technology for UK armed forces and to see the UK defense industry grow stronger through its involvement in the U.S. market. But this vision presupposes a level of bilateral technology cooperation that U.S. leaders appear to be unwilling to concede. (p. 5)

As indicated previously, the UK is clear about the value it places on US military and technological capabilities. However, there is, at times, a UK perception that the US does not appreciate what the UK can bring to the table in this area. As Chao and Niblett note,

In addition, the UK can bring—as it already has brought—valuable technologies to the table for the United States. The UK's track record of useful military technology innovation includes the contemporary examples of the vertical, short take-off and landing engine system and the anti-IED capabilities now deployed in Iraq. A common perception of the threats also means that the UK defense science and technology establishment is focused on solving problems in areas that are of value to the United States, such as counter-terrorism and net enabled warfare. And UK investment in cooperative programs such as the JSF can lessen the development cost for an increasingly-strained U.S. defense budget while decreasing the per unit costs of the system once they go into production. Furthermore, while incomparable in size to the U.S. marketplace, the UK defense market does offer opportunities for major defense contracts. (p. 6)

The question which seems to arise in the minds of UK officials is whether the US believes the UK actually makes a significant contribution to the JSF. This will be a topic addressed by Cranfield researchers as part of the upcoming project. However, it is worthwhile noting that this indirect question (what does the UK think the US is thinking?) is not as valuable as the direct question (what does the US think?), an issue that should be addressed in questions to US officials. At this point, the case could be made that the UK perception is that the US does not place a significant amount of value on what the UK or British industry can bring to the JSF project.

Chao and Niblett note that it should be possible to establish a close bilateral relationship on defence acquisition, as has been developed in far more sensitive areas. Confidence was built up over an extended period of time due to a body of experience, based on established procedures, and a similar long-term process will be needed in the defence acquisition area. There is already an extensive bilateral defence relationship: ten umbrella MOUs between the MoD and the DoD have generated over 100 exchange agreements and 30 project agreements. For the US, this is the largest collaborative relationship. For the UK, it is about 50% of the MoD's Defence Science and Technology Laboratory's projects. And for over a decade, US defence industries have exported over \$1 billion a year to the UK, and UK exports to the US were about \$350 million, with US defence firms having either set up operations in the UK or acquired firms (p. 21). As the authors note:

The heart of any solution must lie in the exceptional closeness of the U.S.-UK political and military relationship. The fact that there are "trusted communities" in the intelligence, nuclear and operational military fields indicates that a similar trusted community could be built in the defense-industrial realm for sharing defense technologies. And, just as with the intelligence and nuclear trusted communities, the answer probably lies in developing a set of special practices, policies, and procedures for defense technology that both sides can have confidence in. (p. 8)

Does It Matter to the US?: The DoD has made clear the damage done to US security policy interests by US export control and technology transfer policies. Under



"Reforming the U.S. Export Control System," the key excerpts from the *Quadrennial Defense Review Report* (2010) clearly note that

Today's export control system is a relic of the Cold War and must be adapted to address current threats. The current system impedes cooperation, technology sharing, and interoperability with allies and partners. It does not allow for adequate enforcement mechanisms to detect export violations, or penalties to deter such abuses. Moreover, our overly complicated system results in significant interagency delays that hinder U.S. industrial competitiveness and cooperation with allies.

The United States has made continuous incremental improvements to its export control system, particularly in adding controls against the proliferation of weapons of mass destruction and their means of delivery. However, the current system is largely out-dated. It was designed when the U.S. economy was largely self-sufficient in developing technologies and when we controlled the manufacture of items from these technologies for national security reasons...

The global economy has changed, with many countries now possessing advanced research, development, and manufacturing capabilities. Moreover, many advanced technologies are no longer predominantly developed for military applications with eventual transition to commercial uses, but follow the exact opposite course. Yet, in the name of controlling the technologies used in the production of advanced conventional weapons, our system continues to place checks on many that are widely available and remains designed to control such items as if Cold War economic and military-to-commercial models continued to apply.

The U.S. export system itself poses a potential national security risk. Its structure is overly complicated, contains too many redundancies, and tries to protect too much. Today's export control system encourages foreign customers to seek foreign suppliers and U.S. companies to seek foreign partners not subject to U.S. export controls. Furthermore, the U.S. government is not adequately focused on protecting those key technologies and items that should be protected and ensuring that potential adversaries do not obtain technical data crucial for the production of sophisticated weapons systems.

These deficiencies can be solved only through fundamental reform. The President has therefore directed a comprehensive review tasked with identifying reforms to enhance U.S. national security, foreign policy, and economic security interests. (pp. 83-84)

The DoD certainly recognizes the new technological landscape. A recent example of the DoD's use of advanced technology is the order by the Air Force for 2,200 Sony Playstation 3 videogame consoles to form the basis of a new super computer (The Economist, 2009).

And recent research indicates that European firms are responding to the difficulties imposed by US policies. Bialos, Fisher, and Koehl (2009) write that "Virtually every interview we conducted highlighted U.S. defense trade controls as a 'barrier' significantly impeding Transatlantic cooperation" (p. 37). Bialos et al. note four key concerns. The first is limits on operational sovereignty, which was noted previously as of particular significance to the UK regarding the JSF. In fact, Bialos et al. writes that "The UK, one of our closest allies, as well as France and Italy, expressed strong concerns about this issue" (p. 113). The other three points noted by Bialos et al. are reliance on US International Traffic in Arms Regulations (ITAR) controlled systems generating risks of schedule delays and increases in



costs; re-export restrictions; and the complications the regulations generate for multinational facilities (p. 113).

Bialos et al. also note, "There is clear evidence, beyond rhetoric, of a behavioral shift in Europe toward 'designing around' or designing out components or subsystems" controlled by the ITAR, "which has a particularly adverse impact on U.S. subsystem and component suppliers" (p. 2). Bialos et al. write,

Over more than a decade, one study after another has highlighted the problems inherent in U.S. export controls—notably the ITAR. While the specifics of these ITAR issues are beyond the scope of this study, the impact of ITAR on the Transatlantic defense market relationship is not. Market participants, U.S. and foreign, consistently report that ITAR slows the speed of obtaining licenses needed for sales and collaboration, limits the release of U.S. technology, creates business uncertainty, and generally makes the process of Transatlantic defense industrial cooperation difficult. Fairly or not, most European governments are concerned about relying on ITAR systems and subsystems because they potentially limit their operational autonomy over major systems (especially in real-time crises), introduce program delays and risks, and curtail their export flexibility for systems with U.S. components.

Years of European talk of "designing around" or "designing out" ITAR have now begun to translate into action, according to market participants—with increased evidence that U.S. ITAR policies and practices, for better or worse, are limiting opportunities for U.S. firms competing in Europe (especially at the subsystem level). This is increasingly true even among our staunchest allies.

The ITAR also inhibits U.S. firms from working with foreign firms on domestic U.S. programs and creates challenges for foreign firms seeking to enter the U.S. market. By declining to release certain information on technologies, the acquisition community can effectively preclude foreign participation.

While strong and well-enforced export controls are an important tool of U.S. national security, it is clear that the U.S. failure to address these concerns will curtail the extent of Transatlantic defense technology sharing, defense cooperation and the development of an open and competitive Transatlantic defense market. (p. 20)

Bialos et al. provide specific examples of this development. On a policy level, the French White Paper "explicitly cites the need for non-ITAR-controlled electronics components to avoid limitations on French freedom of action" (p. 114). In another instance, a country ensured operational sovereignty by "requiring that the program be staffed with domestic engineers free of ITAR restrictions" (p. 114). European firms have developed policies specifically aimed at avoiding the use of ITAR items, developing "dual track" production lines of ITAR and non-ITAR items and favouring suppliers of non-ITAR components (p. 114). Certainly, if ITAR items are superior to the non-ITAR items, the added complications may be worth the added capability. However, if the gap is not that great; ITAR requirements can be a consideration. As noted by Bialos et al., "Where the differential is not great, European governments and firms are increasingly opting for the non-ITAR choice" (p. 114).

Complaints about US export control and technology transfer policy are long-standing, and it is not clear how heavily they will weigh on the decisions of other nations to work with the US on military projects when compared to the costs of cutting-edge military development programs. As Franck et al. (2009) point out, "very few national military establishments can



generate sufficient orders to sustain a weapons source of efficient size in any category" (p. 17). And with the rapid growth of military technology (and the concomitant cost growth) the essential nature of the US in any development program will clearly increase.

However, the US should consider whether it can afford to be indifferent to the willingness of other nations to participate in, and carry some of the costs of, such defence programs. Spreading the cost burden would be presumably appealing to the DoD. And increasing costs also have an impact on the production phase and potential overseas sales. It is an open question whether DoD contracts alone would be sufficient to sustain US military contractors. The Congressional Research Service noted that while the US aviation industry is positioned to compete in the growing global market for civil aircraft, "the extent to which such economic conditions may preserve an adequate U.S. defence industrial base for the development and production of combat aircraft is debatable, however, given the significant differences between civilian and military aircraft requirements and technologies" (Bolkcom, 2009, February 17, p. 17). Even US firms and the DoD may need to focus on overseas sales to sustain programs. And if the US wishes to generate sales to other nations, it will need to address the key issues of operational requirements and sovereignty which have been critical to the UK in the JSF.

Bechat, Rohatyn, Hamre, and Serfaty (2003) note that there will always be a concern in the US, as in all countries, about the leakage of technology as a result of multinational programs. However, there are benefits that also need to be considered:

Governments seeking to strengthen the transatlantic defense relationship must weigh the potential benefits of cooperation against the risks that technology shared with allies might eventually leak to hostile states or subnational groups. Although technology transfer to allies generally involve acceptable risks, those responsible for safeguarding U.S. military technology and for preventing the spread of conventional arms and WMD view technology transfer, even to close allies, as a potential threat and, therefore, as something to be tightly controlled. Close transatlantic cooperation, however, may also reduce the potential risk of technology leakage as it improves each side's perception of sensitive issues and encourages adequate levels of protection. (p. 19)

Bechat et al. note that "there continues to be broad consensus that the U.S. export control system attempts to control too much in light of the widespread diffusion of technologies with defense-related applications" (p. 30). Indeed, one of the key criticisms in the UK that will be explored in the research project is that the US approach is all-encompassing, and does not distinguish between high- and low-tech equipment. There is substantial recognition in the UK that the US has every reason to closely protect cutting-edge military technology, which it developed at significant time and expense. The problem is that the export control and technology transfer policies extend far beyond such sensitive technology to adjacent areas that, to non-American eyes, do not appear to warrant such draconian measures.

However, it is clear that transatlantic cooperation cannot work unless the tech transfer and export control issues are addressed. As Bechat et al. clearly state,

Export control difficulties go to the core of the problems that hamper transatlantic defense cooperation. Changes in the United States and in Europe will be necessary if this critical impediment to enhanced cooperation is to be removed. The U.S. export control system is broken; its technology transfer rules increasingly self-defeating and out of step with broad trends in the global and European economies. Export control reforms in the United States are therefore imperative, including



shrinking the U.S. Munitions List to critical items, instituting greater corporate self-governance with government audits of performance and creating a stronger appeals process for disagreements. (p. 52)

Conclusion

The resurrection of the JSF source code issue is an opportunity to assess the impact of US export control and technology transfer policies, and the UK perspective on this matter is of particular value. The JSF experience may be an indication of structural problems with multinational defence projects, which may be too difficult to be feasible. Technology transfer disputes may be insurmountable. Decisions on the awarding of contracts may be too hard to overcome. Cost increases may be inevitable and excessive and outweigh political and economic interests.

If these problems are inherent in multinational development programmes, perhaps the better option might be for multinational acquisition programs. In such arrangements, there would be fewer states participating in development and more states signing up for purchase of the equipment. This could reduce the impact of some of the problem areas while increasing interoperability. However, there would still be major difficulties, and technology transfer problems, for example, would be reduced, but not eliminated.

But it is not yet apparent that issues such as export controls and technology transfer are an inherently insurmountable barrier to effective programmes. It is certainly important for friends and Allies to appreciate the concerns and interests of the US in protecting its technology. However, the US should acknowledge the views of those who it may wish to participate in a US-led defence programme. And UK views are particularly valuable, particularly as they involve a key programme like JSF. These are the sorts of issues to be addressed as part of the joint NPS-Cranfield research to be conducted over 2010-2011, with the results to be presented at the 2011 symposium.

If multinational defence programmes allow for more efficiency in development and production, and address concerns regarding protection of sensitive technology, they might be a cost-effective option, particularly in an era of tight military budgets and rapid technological advances. However, a prohibitively restrictive export control and technology transfer regime could well negate any benefits from such multinational cooperation, and make it difficult for the US to convince other nations that they should commit themselves to a US-led military development project. In short, if the US does indeed wish to lead multinational programmes that allow for the pooling of finances, talent and technology, this is an issue of importance, not just to the UK and other friends and Allies of the US, but particularly for the US.

References

- Baldwin, T. (2006, December 11). Secret codes clash may sink £140 billion fighter deal. *The Times*. Retrieved December 2, 2008, from <http://www.timesonline.co.uk>
- BBC News. (2006, December 8). *MPs warn over US fighter jet deal*. BBC. Retrieved December 2, 2008, from <http://news.bbc.co.uk/>
- Bechat, J., Rohatyn, F., Hamre, J., & Serfaty, S. (2003, January). *The future of the transatlantic defense community: Final report of the CSIS Commission on Transatlantic Security and Industrial Cooperation in the Twenty-First Century*. Washington, DC: The Center for Strategic and International Studies. Retrieved September 11, 2009, from <http://www.ciaonet.org/book/bej01/>



- Bialos, J., Fisher, C., & Koehl, S. (2009). *Fortresses and icebergs: The evolution of the transatlantic defense market and the implications for U.S. national security policy*. Washington, DC: Center for Transatlantic Relations. Retrieved March 17, 2010, from http://www.acq.osd.mil/ip/docs/fortresses_and_icebergs_vol-one.pdf
- Bolkcom, C. (2009, February 17). *F35 Lightning II Joint Strike Fighter (JSF) program: Background, status, and issues* (Report 7-5700/RL 30563). Washington, DC: Congressional Research Service. Retrieved March 17, 2010, from http://assets.opencrs.com/rpts/RL30563_20090217.pdf
- Bolkcom, C. (2009, February 18). *Proposed termination of Joint Strike Fighter (JSF) F136 alternate engine* (Report 7-5700/RL 33390). Washington, DC: Congressional Research Service. Retrieved March 17, 2010, from <http://www.fas.org/sgp/crs/weapons/RL33390.pdf>
- Cahlink, G. (2006, March 15). U.K. procurement chief warns backup engine dispute threatens JSF deal. *Defense Daily*, 229(50).
- Chao, P., & Niblett, R. (2006, May 26). *Trusted partners: Sharing technology within the U.S.-UK security relationship*. Washington, DC: Center for Strategic and International Studies. Retrieved March 17, 2010, from http://csis.org/files/media/csis/pubs/060526_usukpartnersreport.pdf
- Franck, C., Lewis, I., & Udis, B. (2009, August 9). *Echoes across the pond: Understanding EU-US defence industrial relationships* (Report NPA-GSBPP-09-016). Monterey, CA: Naval Postgraduate School. Retrieved March 17, 2010, from <http://edocs.nps.edu/npspubs/scholarly/TR/2009/NPS-GSBPP-09-016.pdf>
- Gertler, J. (2009, December 22). *F-35 Joint Strike Fighter (JSF) program: Background and issues for Congress* (Report 7-5700/RL 30563). Washington, DC: Congressional Research Service. Retrieved March 17, 2010, from http://assets.opencrs.com/rpts/RL30563_20091222.pdf
- Gray, B. (2009, October). *A review of acquisition for the Secretary of State for Defence: An independent report by Bernard Gray*. Retrieved March 17, 2010, from http://www.aof.mod.uk/aofcontent/downloads/gray/gray_report.pdf
- Haddon-Cave, C. (2009, October 28). *The Nimrod Review*. London: The Stationary Office Limited. HC1025. Retrieved March 17, 2010, from <http://www.official-documents.gov.uk/document/hc0809/hc10/1025/1025.pdf>
- House of Commons Defence Committee. (2005, December 21). *Future carrier and joint combat aircraft programmes* (HC554). Second Report of Session 2005-06. London: The Stationary Office Limited. Retrieved March 17, 2010, from <http://www.publications.parliament.uk/pa/cm200506/cmselect/cmdfence/554/554.pdf>
- House of Commons Defence Committee. (2009, February 26). *Defence equipment 2009* (HC107). Third Report of Session 2008-2009. London: The Stationary Office Limited. Retrieved March 17, 2010, from <http://www.publications.parliament.uk/pa/cm200506/cmselect/cmdfence/554/554.pdf>
- House of Commons Defence Committee. (2010, March 4). *Defence equipment 2010* (HC99). Sixth Report of Session 2009-2010. London: The Stationary Office Limited. Retrieved March 17, 2010, from <http://www.publications.parliament.uk/pa/cm200910/cmselect/cmdfence/99/99.pdf>
- House of Commons Foreign Affairs Committee. (2010, March 28). *Global security: UK-U.S. relations* (HC 114). Sixth Report of Session 2009-2010. London: The Stationary Office Limited. Retrieved March 29, 2010, from <http://www.publications.parliament.uk/pa/cm200910/cmselect/cmfaaff/114/114.pdf>
- Inspector General. (2008, March 6). *Security controls over Joint Strike Fighter classified technology* (Report No. D-2008-058). Arlington, VA: DoD Inspector General.



- Retrieved March 17, 2010, from <http://pogoarchives.org/m/ns/dod-ig-report-20080306.pdf>
- Joint Strike Fighter Program Office. (2007). *Annex A (April 2007 Revision) estimated JSF air vehicle procurement quantities*. Retrieved December 2, 2008, from http://www.jsf.mil/downloads/down_documentation.htm
- Joint Strike Fighter Program Office. (2009, March 19). *United Kingdom announces F-35 Joint Strike Fighter purchase*. Retrieved April 3, 2009, from <http://www.jsf.mil/news/news2009.htm>
- La Franchi, P. (2006, July 4). Australia demands JSF resolution. *Flight International*. Retrieved March 15, 2010, from <http://www.flightglobal.com>
- McGhie, T., & Gee, J. (2006, February 26). Anglo-US defence deals in jeopardy. *Financial Mail*, p. 1.
- Metz, P. (2005). The Joint Strike Fighter—An international enterprise. *Flight Test—Sharing Knowledge and Experience* (pp. KN-3-1 -- KN3-8). Meeting Proceedings RTO-MP-SCI-162, Paper KN3. Retrieved March 17, 2010, from [http://ftp.rta.nato.int/public//PubFullText/RTO/MP/RTO-MP-SCI-162///MP-SCI-162-\\$KN3.pdf](http://ftp.rta.nato.int/public//PubFullText/RTO/MP/RTO-MP-SCI-162///MP-SCI-162-$KN3.pdf)
- Ministry of Defence. (2005, December). *Defence industrial strategy* (Cm6697). Defence White Paper. London: The Stationary Office Limited.
- Ministry of Defence. (2010, January 12). *Strategic trends programme: Global strategic trends—out to 2040* (4th ed.). Retrieved February 26, 2010, from <http://www.mod.uk/DefenceInternet/MicroSite/DCDC/OurPublications/StrategicTrends+Programme/>
- Ministry of Defence. (2010, February 12). *Strategic trends programme: Future character of conflict*. Retrieved February 26, 2010, from <http://www.mod.uk/DefenceInternet/MicroSite/DCDC/OurPublications/Concepts/>
- Ministry of Defence. (2010a, February). *Adaptability and partnership: Issues for the Strategic Defence Review* (Cm7794). London: The Stationary Office Limited.
- Ministry of Defence. (2010b, February). *The defence strategy for acquisition reform* (Cm7796). London: The Stationary Office Limited.
- National Audit Office. (2008, December 18). *Ministry of Defence: Major projects report 2008—project summary sheets* (HC64-II). Session 2008-2009. London: The Stationary Office Limited.
- Norton-Taylor, R. (2010, January 12). MoD to slash jet fighter orders as it struggles to save aircraft programme. *The Guardian*. Retrieved March 25, 2010, from <http://www.guardian.co.uk>
- O'Connell, D. (2004, July 25). Britain fights for larger stake in JSF. *The Times*. Retrieved December 2, 2008, from <http://business.timesonline.co.uk/>
- O'Rourke, R. (2009, September 16). *F-35 Joint Strike Fighter (JSF) program: Background and issues for Congress* (Report 7-5700/RL 30563). Congressional Research Service. Retrieved March 17, 2010, from http://assets.opencrs.com/rpts/RL30563_20090916.pdf
- Scott, R. (2009, January). Preparing for take-off: UK ramps up JSF carrier integration effort. *Jane's International Defence Review*, 44-49.
- The Economist. (2009, December 12). The military-consumer complex. *The Economist*, 393(8661), 16.
- US Congress. (2006, October 17). *National defense authorization act for FY2007* (HR 5122/PL 109-364). Retrieved March 17, 2010, from <http://www.govtrack.us/congress/bill.xpd?bill=h109->



US DoD. (2010, February). *Quadrennial defense review report*. Retrieved March 11, 2010, from http://www.defense.gov/QDR/images/QDR_as_of_12Feb10_1000.pdf

Warwick, G. (2006, March 16). UK and Australia threaten to pull out of Lockheed Martin JSF over tech transfer delays. *Flight International*. Retrieved March 5, 2010, from <http://www.flightglobal.com>



Global Cooperation and Competition in the Defense and Aerospace Industries

Raymond Franck—Raymond (Chip) Franck, PhD, Senior Lecturer, Graduate School of Business & Public Policy, Naval Postgraduate School, retired from the Air Force in 2000 in the grade of Brigadier General after 33 years of commissioned service. He served in a number of operational tours as a bomber pilot; staff positions, including the Office of Secretary of Defense and Headquarters, Strategic Air Command; and as Professor and Head, Department of Economics and Geography at the US Air Force Academy. His institutional responsibilities at NPS have included the interim chairmanship of the newly formed Systems Engineering Department (July 2002 to September 2004), serving as Associate Dean for Academic Operations (December 2007 to present), teaching a variety of economics courses, and serving on a number of committees to revise curricula for both the Management and Systems Engineering disciplines. His research agenda focuses on defense acquisition practices and military innovation.

Ira Lewis, PhD, is Associate Professor of Logistics, Graduate School of Business and Public Policy, Naval Postgraduate School, Monterey, CA. His interests include transportation, public policy, and the international defense industry.

Bernard Udis, PhD, is Professor Emeritus of Economics at the University of Colorado at Boulder and Visiting Research Professor at the US Naval Postgraduate School. He has also served as Distinguished Visiting Professor of Economics at the US Air Force Academy and as a William C. Foster Fellow at the US Arms Control & Disarmament Agency. His NATO Research Fellowship examined the costs and benefits of offsets in defense trade.

Professor Udis' published work includes three books: *The Economic Consequences of Reduced Military Spending* (editor, 1973), *From Guns to Butter: Technology Organizations and Reduced Military Spending in Western Europe* (1978), and *The Challenge to European Industrial Policy: Impacts of Redirected Military Spending* (1987). In addition, he has published numerous articles in scholarly journals on defense industries and military power. These include *Offsets as Industrial Policy: Lessons From Aerospace* (with Keith Maskus, 1992), and *New Challenges to Arms Export Control: Whither Wassenaar?* (with Ron Smith, 2001). A number of his works are considered classics in defense economics and have been reprinted in collections such as *The Economics of Defence* (2001) and *Arms Trade, Security and Conflict* (2003).

Professor Udis' current research focuses on competition and cooperation in the aerospace industries of the US and the EU.

Executive Summary

This is a summary of the report cited above for inclusion in the Naval Postgraduate School's Proceedings of the Seventh Annual Acquisition Research Symposium (May 2010). The report itself greatly exceeds the length guidelines for the Proceedings. The topics raised here are discussed in greater detail within the body of that report.

This is the third report from our project to better understand the evolving international-defense industrial base. We've pursued two basic aims:



- First, to better “map the terrain” in the international defense marketplace, which is steadily becoming more complex; and
- Second, to consider the utility of various perspectives (or paradigms) in understanding the forces driving the changes in that industrial base.

As a starting point in this multiyear effort, we undertook to study the interactions between the US and Europe (primarily European Union (EU) and North Atlantic Treaty Organization (NATO) members). Well into this project, we’re still considering primarily transatlantic defense industrial affairs.

OUR CASES

1. THE KC-X COMPETITION

We continue our inquiry into the KC-X saga, which has consumed pretty much the entire first decade of the 21st century. It has also been a major theme in our ongoing research project. In previous reports, we’ve considered the EADS strategy to enter the North American defense market (the KC-30 candidate for KC-X being a major part of that strategy). In our second report, we considered the KC-X (or KC-45) competition between EADS (with its partner Northrop Grumman, NG) and Boeing, covering the major events of the competition from the original Request for Proposal (RFP) through the contract award to EADS-NG in February 2008, which included the Boeing protest, the GAO’s decision sustaining that protest, and the DoD’s abortive attempt to quickly re-compete the project. Among other things, we concluded that the government side of the affair was better described by “quarrelsome committee” than by the traditional notion of the “sovereign monopolist.” In this report, we give Graham Allison’s Model III (governmental politics) a preliminary test-drive as a paradigm for understanding the operations of the quarrelsome committee that has presided over the KC-X Affair.

Overall, we see a confluence of trends with the KC-X Affair. The growth in the size, complexity, and expense associated with new systems, combined with the post-Cold War defense industrial consolidations, mean fewer serious bidders for new projects. When the US Air Force needed a new tanker, it found only one US supplier: Boeing. However, EADS, which offered the KC-30 and had partnered with Northrop Grumman (NG), provided strong competition. Although the EADS-NG team offered an attractive package that included substantial US industrial participation, Boeing’s political supporters were nonetheless able to appeal to nationalist sentiment against a “foreign” supplier. The results so far resemble a political quagmire. We also think that the “quarrelsome committee” (a dysfunctional example of governmental processes) is now a serious methodological challenger to the “sovereign monopsonist” model.

2. AIRBUS A400M DEVELOPMENT

Our second case considers the difficulties that have attended the development of the Airbus A400M military transport aircraft—primarily the engines. Among other things, it’s an interesting complement to our previous research about the difficulties encountered in the development of the Boeing 787 “Dreamliner” commercial transport aircraft. As we’ve noted in previous reports, the scale and expense of major weapons systems frequently exceeds the capacity of firms, and also nation-states. Among other things, this means that technical complexity has been accompanied by management complexity—at a number of levels. These include management of technology (or systems engineering) of integrating complex systems and of systems-of-systems. They also include the management of complex systems of partnerships among coalitions of firms and defense establishments. The Boeing



787 experienced difficulties in the management of an extensively outsourced development project and a complex supply chain. In the case of the A400M, the political choices made by the nations (and sponsors) drove the choice of a local engine—developed by a pickup team of European firms. The result was extensive delays that have been costly in a number of ways. Basically, the political imperatives associated with managing the A400M manifested themselves in the enormous technical difficulties encountered in developing those engines.

A main theme in the case of the A400M is the interaction of globalization, sovereignty, and complexity. The European participants chose an EU-based defense consortium as a method of reconciling globalization and “sovereignty” (in a regional sense). The geographic and political logic of that consortium led to European engine development, despite an alternative that was much closer to fruition (and with much lower risk) from North America. The result has been very bad for Airbus and its prospective customers—extensive delays and cost overruns. With those adverse developments, A400M buyers have either dropped away, hedged their bets, or become visibly less enthusiastic about pouring still more resources into the project. While the A400M project has recently been refinanced and revived, we suspect the A400M story will, in the end, more closely resemble the Concorde’s than the A320’s.

3. The (Maybe) Emerging Nordic Defense Bloc

Finally, we undertake (we think) a fresh look at the maybe-emerging Nordic defense bloc. Nordic powers are making a number of serious and successful efforts to enhance cooperation among their militaries—with varying degrees of success. But, will the natural attractions of geography, common culture, and common interests overcome the simultaneous attractions of nationalism and outside partners? This tension has, *inter alia*, manifested itself in the competition between the Swedish Viggen and the American Joint Strike Fighter as the next-generation Multi-role Combat Aircraft (MRCA). In our research, we’ve been fortunate to have received generous cooperation from a number of international respondents (promised anonymity). They’ve provided valuable, inside information about both the attractive and centrifugal forces that affect the Nordic players. Whether the Nordic group turns out to be a nascent defense bloc, or just a dead end, we believe the directions that Nordic defense establishments take will have significant implications for defense industrial affairs on both sides of the Atlantic.

Our discussion of the Nordic bloc is an excellent example of the issues associated with regional partnerships. Regional partnerships are attractive in principle and should be especially attractive given the many commonalities among the Nordic countries. In practice, however, there are many “devils in the details.” Indeed, some of those details surfaced in our interviews with knowledgeable officials from those Nordic countries.

Overall: Taken as a whole, we think this particular report has furthered our research agenda of better understanding the nature of the contemporary defense industrial base—in all its global complexities—and in understanding some of the more likely future developments. We also think we’ve surfaced a fresh mode of analysis (Allison’s Model III) as a candidate vehicle for better understanding the dynamics of contemporary defense industrial affairs.



Panel #19 – System-of-Systems Design & Development

Thursday, May 13, 2010

11:15 a.m. –
12:45 p.m.

Chair: Joseph L. Yakovac, Jr., LTG, USA, (Ret.), Naval Postgraduate School; former Military Deputy to the Assistant Secretary of the Army (Acquisition, Logistics & Technology)

Acquisition Risks in a World of Joint Capabilities

Maureen Brown and Anita Raja, The University of North Carolina at Charlotte, and Robert Flowe, Office of the Deputy Under Secretary of Defense (Acquisition & Technology)

A Technique for Evaluating Complex System of Systems Designs

Stephen Blanchette Jr., Carnegie Mellon University and **Steven Crosson**, PEO Integration

System Development and Risk Propagation in Systems-of-Systems

Muharrem Mane and Daniel DeLaurentis, Purdue University



Acquisition Risks in a World of Joint Capabilities

Maureen Brown—Mary Maureen Brown is a Professor of Public Administration at the University of North Carolina at Charlotte. Her research interests center on systems engineering, inter-organizational theory, research methods, and program planning.

Mary Maureen Brown, PhD
Department of Political Science
University of North Carolina Charlotte
9201 University City Blvd
Charlotte, NC 28223-0001

Robert Flowe—Robert Flowe serves on the AV SOA Functional Team, OSD.

Robert Flowe
AV SOA Functional Team
OUSD(AT&L)/ARA/EI
CG-2 Suite 900

Anita Raja—Anita Raja is an Associate Professor of Software and Information Systems at the University of North Carolina at Charlotte. She received a PhD in Computer Science from the University of Massachusetts Amherst in 2003. Her research interests are in the field of artificial intelligence, including distributed decision-making, meta-cognition and bounded-rationality.

Anita Raja, PhD
Department of Software and Information Systems
University of North Carolina at Charlotte
9201 University City Blvd
Charlotte, NC 28223-0001

Abstract

This study reports some preliminary results of a research agenda that seeks to address the absence of tested metrics to provide early indication of the acquisition risks of interdependent programs. The overall goal of this research is to forge new ground on uncovering early indicators of interdependency acquisition risk, so appropriate governance mechanisms can then be isolated. Funding will allow the ability to 1) expand on an existing database of Major Defense Acquisition Program (MDAPs) performance data, 2) analyze the MDAPs to characterize the risks attributable to interdependence, and 3) determine whether acquisition setbacks cascade to downstream interdependent MDAP programs. The deliverables of the effort are 1) a code book of DoD data acquisition items that can be employed in future research efforts, and 2) the results of the investigation of the cascading risks of interdependent acquisition efforts.

In short, preliminary results indicate that perceptions of risk may prove influential on downstream program performance. In terms of the direct influence that an upstream program's performance might exert on downstream program performance, weak, but statistically significant, relationships were noted in three areas. The next steps of the research are to 1) expand the dataset to include FY 2009 data, 2) document acquisition data, 3) collect a number of indicators on program interdependency, and 4) test a number of interdependency diversity metrics in terms of their ability to provide insights on program performance.



Introduction

This research effort addresses two critical problems: 1) Data on major defense acquisition programs is piecemeal and fragmented, thus hindering acquisition research; 2) There is an absence of tested metrics to provide early indication of the acquisition risks of interdependent programs.

A wealth of research indicates that the use of programmatic networks in the public sector is clearly on the rise (Weber & Khademian, 2008). Noting that firms do not act by themselves, Granovetter (1973) witnessed that organizations are deeply embedded in “networks of external relationships” that influence the exchange of resources and capabilities among them. Increasingly, state-, local-, and federal-level agencies are turning to joint interdependent programs to address gaps that only cross collaborative initiatives can span. Yet, as discussed below, the study of interdependency and its effects on program performance have yielded too few tangible results. For purposes of the discussion below, jointness, interdependency, exchange, and partnerships, all refer to a similar concept: the notion that autonomous organizations build relationships to provide capabilities that, when looked at in totality, form network structures. Additionally, at the individual pair-wise level, the exchanges are manifested as explicit transactions.

Scholars have long contended that many contemporary policy challenges, and their associated solutions, lie across organizational domains outside the jurisdiction of any one agency (Gage & Mandell, 1990; Alexander, 1995; Agranoff, 2003). Milward and Provan (2001) show that public policy arenas are inherently crosscutting; the requisite knowledge and corresponding solutions are not localized, but are instead distributed across a range of agencies and organizations. The DoD’s transformation to joint capabilities is in keeping with the ongoing trends. Historically, acquisition investments at the DoD had been proposed as individual materiel solutions, typically championed by the armed service for which the product was to be obtained. This gave rise to discrete systems designed in accordance with the individual service requirements. When called upon to operate in a joint, multi-service environment, these systems exhibited problems interacting effectively with other service systems.

The Transformation to Joint Capabilities attempts to provide military forces with the capability to adapt quickly to new challenges and unexpected circumstances by leveraging a wider range of assets. Central to the Transformation was the desire for enhanced coordination among agencies and across all levels of government (coalition, federal, state, and local). In addressing the need for interagency cooperation, Vice Chairman of the Joint Chiefs of Staff Admiral Giambastiani (2004) claimed that the integrated force had to become interdependent. That is, it must be capabilities-based, collaborative, and network centric. Military efforts require the ability to conduct high-level, or large-scale, vertical and horizontal collaboration. That means up and down the chain of command and across all capabilities and forces.

While DoD agencies are expected to embrace joint capabilities, literature findings regarding the risks and best practice mechanisms of joint interdependent activities lag far behind. Whereas early research did provide some insights, the research activities have stalled and progress is lacking. For example, back in 1937, Coase found that interdependencies are based on mutual exchanges that can be examined at the transaction level. He argued that these transactions accrued costs that could be attributed to establishing the rules of engagement, enforcing agreements, and monitoring compliance. Unfortunately, specific cost functions were never isolated.



In 1967, Thompson contributed to the research by offering a tripartite model that focused on the configurations of the transaction flows. Sequential flows involved handoffs between partners. Pooled flows involved partners that drew down from a common source of assets and the flows of reciprocal relationships involved feedback mechanisms. Much of the research to date has been based on the anecdotal findings of small *n* case studies (Isett & Provan, 2005; Meier & O'Toole, 2008). While the three configurations provide a starting point for understanding interdependent activities, the reality of today's activities are far more complex. In short, it is not unusual for the acquisition or production of a service to incorporate multiple configurations with resources flowing in and out across organizations of public and private entities. As such, the most common configuration is the "mixed" pattern incorporating all three of the configurations and a wide array of nodes, assets, channels and zones.

Exchange theorists argue that organizations develop interdependent relationships with other organizational entities to either obtain critical resources or provide critical capabilities. They also assert that interdependent relationships exhibit high levels of uncertainty due to participant constraints (Miles & Snow, 1978).

Shirking or defection of a network member can have dire consequences on the survival and performance of the network in total and network participants in general. Because of the nature and influence of the ties or interdependencies that bind organizations, Levinthal's (1997) research indicated that increasing the density of the interdependencies that connect the organizations affects the complexity of the "landscape" in which it operates. Levinthal (1997) finds that these interconnections or flows can yield nonlinear consequences that often involve multiplier effects based on the nature of the interdependencies in the system.

Apparently, the value chain of the joint capabilities is laden with junctions and bifurcations where delay, defection, or shirking can occur. In fact, in 1999, Rosen argued that the uncertainty that arises from a relationship is the definition of "complexity." And that "complexity" can only be understood by examining the links that bind. If Rosen is correct, then DoD acquisition is reaching unprecedented complexity. A network analysis of MDAP data interdependencies suggests overwhelming complexity (see Figure 1). The 78 MDAPs in 2005 demonstrated 989 unique links to other MDAPs as well as non-MDAP programs. The yellow links indicate medium risk relationships and the red links show high risk links.

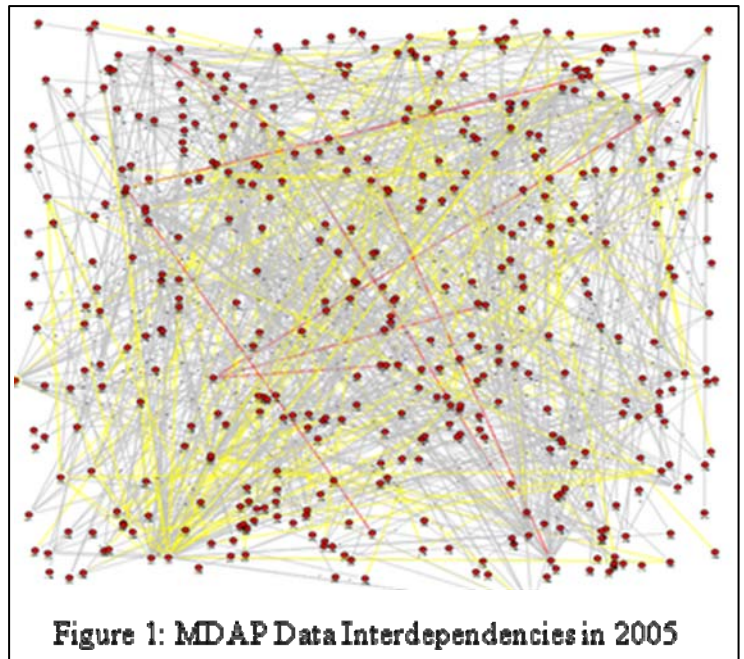


Figure 1: MDAP Data Interdependencies in 2005

Despite the activities, ten years ago Agranoff and McGuire (2001) wrote that "there are many more questions than answers in network management" and the assertion continues to ring true. Apparently, the field is rich in anecdotal findings but poor in empirical evidence (Alexander, 1995). Oliver and Ebers (1998) liken the state of the field as a messy situation marked by a cacophony of heterogeneous concepts, theories, and research results. While the growth is clearly on the rise, DoD acquisition is moving forward with little

insights into the risks and threats of joint efforts. Without a deep understanding of the risks and threats that interdependent efforts encounter, governance mechanisms that can help to insure acquisition success are beyond reach. Given the pace at which joint efforts are pursued, early indicators of acquisition risk are needed to help isolate the critical governance mechanisms that will mitigate performance shortfalls.

Additionally, of utmost concern is the state of current acquisition data. While several initiatives are underway to compile the sundry datasets the fragmented acquisition data puts decision-making at risk for Type I and Type II errors (false positive and false negative, respectively). Funding for this research effort will allow the ability to integrate, cleanse, and normalize authoritative datasets for the purpose of advanced research. It will also provide the ability to document the acquisition data for future research purposes. The documentation will allow the research community to forge new insights in the acquisition arena.

The study of DoD acquisition efforts for gaining insights on interdependency is especially fruitful. First, the DoD has rich, but fragmented and piecemeal, datasets on some of the important key interdependencies of major defense acquisition programs. Second, the movement toward joint capabilities makes an understanding of interdependencies especially critical. Finally, given the frequency with which government agencies are moving toward joint initiatives, findings based on DoD programs may prove instrumental to a wide range audience. The research below examines the ties that bind organizations in light of two different types of transactions: data ties and funding ties. And it begs the questions: do the ties result in cascading acquisition risks? The research will act as a catalyst for a long-term research program on the risks and governance mechanisms of joint acquisition initiatives.

Research Objectives: The objectives of the research are as follows:

1. To map program interdependence to reveal the directionality of influence (i.e., “upstream”/“downstream”) of cause-effect relationships,
2. To test the cascading risks that upstream programs exert on downstream programs in light of data and funding exchanges,
3. To test the extent to which the cost overruns of upstream programs cascade on to interdependent downstream programs,
4. To test the extent to which schedule delays cascade on to interdependent downstream programs,
5. To use the findings to make recommendations on potential governance mechanisms that may prove capable of mitigating the risks of interdependencies, and
6. To provide a research code book of acquisition data elements for future research efforts.

The remaining discussion provides an interim report on some of the findings to date, based on the interdependent activities identified in the Defense Acquisition Executive Reviews (DAES). The findings below examine the influence of interdependencies on a number of program performance measures.

Research Methods: Because of data availability issues, the unit of analysis for this research was restricted to Major Defense Acquisition Programs (MDAPs). Starting in the 2005 time period, Major Defense Acquisition Program managers were asked to provide reports on what they considered to be the critical interdependencies of their given program.



They identified upstream and downstream connections and indicated the perceived risk level (red, yellow, green). Because of a small number of red risks, the risk variable was recoded to reflect “no-risk,” “risky.” Hence the “risk” variable is binomial in nature. The research findings described below are based on the influences these interdependencies might exert on program performance.

Program performance is considered from multiple vantage points. Table 1 provides all the variables used in this research and the documents they were derived from. In short, performance is based on 1) annual cost variance figures (for total cost variance and the subsets of schedule, estimation, and engineering cost variance in millions), 2) DAES breaches (schedule, performance, and Research, Development, Testing, and Evaluation (RDT&E) breaches), and 3) percent cost growth from the original RDT&E estimates. All variables were derived from Selected Annual Reports (SARs) and DAES reports. Because 2005 marked the first year that the DAES reports began reporting interdependencies, the analysis reported below is based on MDAP performance in fiscal years 2005, 2006, and 2007 (note: SAR reports were not reported in FY2008 due to the new presidential administration).

Table 1. Variables Used in the Research

Unit of Analysis: Major Defense Acquisition Programs	
Variable	Data Source
Count of APB Schedule Breaches Count of APB Performance Breaches Count of APB RDT&E Breaches Count of APB PAUC Breaches	DAES
Total Cost Variance Engineering Cost Variance Schedule Cost Variance Estimation Cost Variance	SAR
Percent Cost Growth	SAR
Data Exchange Interdependencies	DAES Interdependency Charts

As identified above, MDAP program managers were asked to provide insight on what they perceived to be the program’s interdependencies. They also reported on the direction of the interdependency (inbound, outbound, and bidirectional) and the risk of the interdependency. The risk is based on the sender’s perceived risk with a downstream receiver. Because performance data on non-MDAP programs were not available for analysis, the findings considered only the interdependencies that existed among MDAPs.

Findings: Two major sets of findings are discussed below. The first is based on the influence of the sender’s perceived risk with the downstream receiver’s performance (during years 2005-2007). The sender’s “perception of risk” is also considered in light of their own individual performance (during years 2005-2007). To measure the sender’s perceived risk on their individual performance, the mean of the risk of all the relationships was calculated to provide an overall “risk” level for each MDAP. The second set of findings is based on the sender’s performance and its influence on the receiver’s performance. Table 2 provides the means and standard deviations of all variables used in the research. The number of

downstream programs per upstream MDAP ranged from 1 to 23, with a mean of 5. A total of 873 relationships were analyzed over the three-year time period.

Table 2. Descriptive Statistics of Variables Employed in the Research

	N	Mean	Std. Deviation
Upstream APB Schedule Breaches	269	.36	.481
Upstream APB RDT&E Breaches	522	.09	.289
Upstream APB PAUC Breaches	522	.08	.266
Upstream APB Performance Breaches	522	.13	.333
Downstream APB Schedule Breaches	518	.27	.444
Downstream APB Performance Breaches	522	.06	.240
Downstream APB RDT&E Breaches	522	.11	.319
Downstream APB PAUC Breaches	522	.11	.312
Upstream Perceived Risk	522	1.1073	.30977
	N	Mean	Std. Deviation
Upstream Percent Cost Growth	840	.08	1.00
Upstream Engineering Cost Variance	711	5.53	91.06
Upstream Schedule Cost Variance	711	4.76	38.31
Upstream Estimation Cost Variance	711	.32	159.33
Upstream Total Cost Variance	840	4.75	208.20
Downstream Percent Cost Growth	446	.014	.10
Downstream Engineering Cost Variance	351	-6.12	219.25
Downstream Schedule Cost Variance	351	3.77	25.05
Downstream Estimation Cost Variance	351	-7.95	324.36
Downstream Total Cost Variance	449	-38.79	400.26

Interdependency Risk. The first set of tests sought to determine if the sender's perceived risk of the relationship influenced its partnering receiver. In terms of the "partner risk" variable, xx% of the MDAP program managers identified no risk in the partnerships. Xx% indicated some degree of risk in the relationships. Of the 56 programs that indicated some risk, the risk ranged from a low of 1.1 to a high of 2 (recall that the variable ranges from "1" to "2," with "1" indicating no risk). Correlation coefficients were then obtained (see Table 3). The data show that the Manager's Perception of Risk is *negatively* with the partner's total cost variance, engineering cost variance, and estimation cost variance. Interestingly, risk was correlated with the downstream partner's performance, RDT&E, and PAUC breaches, but in a *positive* direction. The fact that an upstream partner's perception of risk might result in an *increase* in the number of DAES' breaches illustrates the detrimental influence that upstream programs might exert on their downstream partners. But, why the *negative* relationship with cost variance? Why would an increasing perception of risk on the upstream program result in reducing the cost variance of its downstream partners? Perhaps the answer lies in the old adage that perceptions of risk result in increased attention. Perhaps, under high-risk situations, program managers are more cognizant of the risk and act to mitigate the detrimental effects. More research is clearly warranted to tease out why these correlations are demonstrated.

Table 3. Bivariate Correlation with MDAP Program Manager's Perception of Risk

MDAP's Percent Cost Growth	-.024
MDAP's Total Cost Variance	.019
MDAP's Engineering Cost Variance	-.081*
MDAP's Schedule Cost Variance	-.007
MDAP's Estimation Cost Variance	.05
MDAP's Schedule Breaches	.05
MDAP's Performance Breaches	.11*
MDAP's RDT&E Breaches	.13**
MDAP's PAUC Breaches	-.05
Downstream MDAP's Percent Cost Growth	-.07
Downstream MDAP's Total Cost Variance	-.12**
Downstream MDAP's Engineering Cost Variance	-.22**
Downstream MDAP's Schedule Cost Variance	-.05
Downstream MDAP's Estimation Cost Variance	-.11*
Downstream MDAP's Schedule Breaches	.07
Downstream MDAP's Performance Breaches	.13**
Downstream MDAP's RDT&E Breaches	.09*
Downstream MDAP's PAUC Breaches	.12**
* p < .05 ** p < .00	

The next step of the “risk” analysis sought to isolate whether the MDAP’s mean risk score influenced their own specific performance. The results show that the manager’s perception of risk is positively correlated with the program’s engineering cost variance. Outside of engineering cost variance, despite recognition of risk, no notable correlations were demonstrated in terms of DAES breaches, cost variance, or cost growth.

The second set of tests sought to identify whether the performance of upstream programs exert influence on their downstream partner’s performance. Because an upstream program’s influence would not expect to have an immediate effect, the data were lagged a year. In other words, one might expect that the negative effects of an upstream program would influence their downstream partners one year out. Table 4 shows the results of the sender’s performance on the downstream partners. The results show little influence on APB breaches. A weak, but statistically significant, relationship was demonstrated between the number of upstream program RDT&E breaches and the number of downstream program RDT&E breaches. In terms of the influence of the upstream program’s percent growth, it showed no correlation with the downstream program’s percent growth. Two of the cost variance relationships also showed weak, but statistically significant, correlations. The sender’s total cost variance appeared to exert some influence on the downstream program’s schedule variance. In addition, as the sender’s engineering cost variance rose, a subsequent rise was also noted in the downstream program’s percent cost growth.

**Table 4. The Influence of the Sender’s Performance on Downstream Programs
FY 2005-2007 (Lag of One Year)**

		SENDER MDAP								
		APB Schedule Breaches	APB RDT&E Breaches	APB Performance Breaches	APB PAUC Breaches	Total Cost Variance	Engineering Cost Variance	Schedule Cost Variance	Estimation Cost Variance	Percent Cost Growth
D O W N S T R E A M P R O G R A M	APB Schedule	-.03	.04	-.01	.05					
	APB RDT&E	-.00	.07*	.02	-.05					
	APB Performance	.06	.04	-.00	.01					
	APB PAUC	-.03	.05	-.02	.00					
	Total Cost Variance					-.02	.09	-.06	.01	-.01
	Engineering Cost Variance					.02	.09	-.06	.02	-.01
	Schedule Cost Variance					.09*	-.04	.05	-.02	-.00
	Estimation Cost Variance					-.08	.09	.02	.03	.02
	Percent Cost Growth					-.02	.12**	-.00	.02	-.00

Conclusions

The results discussed above are the preliminary interim results of a segment of research that seeks to identify the influences that interdependencies might exert on

acquisition program performance. In short, the results indicate that perceptions of risk may prove influential on downstream program performance. In terms of the direct influence that an upstream program's performance might exert on downstream program performance, weak, but statistically significant, relationships were noted in three areas. In subsequent months, we will complete the data collection effort and construct and test a series of interdependency metrics on program performance. The data will be modeled using traditional statistical approaches to assess causality. Additionally, we will employ Markov Decision Process (MDP)-based methods (Puterman, 1994) to take into account the cost and schedule variance specifications from the n-ordered downstream programs and produce a specification of the best possible budget trimming options for the decision-maker. Formally, a MDP is a probabilistic model of a sequential decision problem, in which states can be perceived exactly, and the current state and action selected determine a probability distribution on future states (Bertsekas, 1987). Specifically, the outcome of applying an action to a state depends only on the current action and state (and not on preceding actions or states). We assume that state changes in our model occur only at discrete instances of time, allowing us to model the network as a discrete event dynamic system (DEDS) and plan to employ MDPs. Our model facilitates the data acquisition process since we iteratively refine the state features critical to the decision-making. The action space will capture information about the funders, including changes in level of funding. We plan to model the probability of transitions from one state to another empirically by using existing data. The "Reward" function will be the presence of a schedule delay and cost variance that occurs in n-ordered downstream programs. Hence, we will be assessing various interdependency metrics in light of statistical and MDP methods to isolate the most feasible method for understanding interdependencies.

In short, the next steps of the research are to 1) expand the dataset to include FY 2009 data, 2) document existing acquisition data sources, 3) collect a number of indicators on program interdependency, and 4) test a number of interdependency diversity metrics in terms of their ability to provide insights on program performance.

References

- Agranoff, R. (2003, October). *A new look at the value-adding functions of intergovernmental networks*. Paper presented at the National Public Management Research Conference, Georgetown University, Washington, DC.
- Agranoff, R., & McGuire, M. (2001). Big questions in public network management research. *Journal of Public Administration Research and Theory*, 113, 295-396.
- Bertsekas, D. (1987). *Dynamic programming: Deterministic and stochastic models*. Upper Saddle River, NJ: Prentice Hall.
- Coase, R. (1937). The nature of the firm. *Economica*, 4(16), 386–405.
- Gladwell, M. (2002). *The tipping point: How little things can make a big difference*. Gladwell Publishers.
- Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360-1380.
- Isett K., & Provan, K. (2005). The evolution of dyadic interorganizational relationships in a network of publicly funded nonprofit agencies. *Journal of Public Administration Research and Theory*, 15, 149-165.
- Joint Vision, 2020. (2009). Retrieved April 1, 2009, from http://www.fs.fed.us/fire/doctrine/genesis_and_evolution/source_materials/joint_vision_2020.pdf



- Jones, C., Hesterly, W.S., & Borgatti, S.P. (1997). A general theory of network governance: Exchange conditions and social mechanisms. *Academy of Management Review*, 22(4), 911-945.
- Levinthal, D., & Warglien, M. (1999). Landscape design: Designing for local action in complex worlds. *Organization Science*, 10(3), 342-357.
- Levinthal, D. (1997). Adaptation on rugged landscapes. *Management Science*, 43(7), 934-950.
- Mandell, M.P. (Ed.). (2001). *Getting results through collaboration: Networks and network structures for public policy and management*. Westport, CT: Quorum Books.
- Meier, K.J., & O'Toole, L.J. (2001). Managerial strategies and behavior in networks: A model with evidence from u.s. public education. *Journal of Public Administration Research and Theory*, 11, 271-295.
- Meier, K.J., & O'Toole, L.J., Jr. (2008). Management theory and Occam's Razor: How public organizations buffer the environment. *Administration & Society*, 39(8), 931-958.
- Miles, R.E., & Snow, C.C. (1978). *Organizational strategy, structure, and process*. New York: McGraw-Hill.
- Miles, R.E., & Snow, C. (1992). Causes of failure in network organizations. *California Management Review*. Summer.
- Milward, H.B., & Provan, K. (2001, July/August). do networks really work? a framework for evaluating public sector organizational networks. *Public Administration Review*, 61(4), 414-423.
- Milward, H.B., & Provan, K.G. (1998). Measuring network structure. *Public Administration*, 76, 387-407.
- Oliver, A., & Ebers, M. (1998). Networking network studies: An integrative framing of conceptual dimensions in the study of interorganizational relationships. *Organization Studies*, 19(4), 549-583.
- O'Toole, L.J., Jr. (1997). Treating networks seriously: Practical and research-based agendas in public administration. *Public Administration Review*, 57(1), 45-52.
- O'Toole, L.J., Jr., & Meier, K.J. (2004). Desperately seeking Selznick: Cooptation and the dark side of public management in networks. *Public Administration Review*, 64(6), 681-693.
- Podolny, J.M., & Page, K.L. (1998). Network forms of organization. *Annual Review of Sociology*, 24, 57-76.
- Powell, W.W. (1990). Neither market nor hierarchy: Network forms of organization. In B.M. Staw & L.L. Cummings (Eds.), *Research in organizational behavior* (pp. 295-336). JAI.
- Provan, K.G., & Kenis, P.N. (2008). Modes of network governance: Structure, management, and effectiveness. *Journal of Public Administration Research and Theory*, 18(2), 229-252.
- Provan, K.G., & Milward, H.B. (2001). Do networks really work? A framework for evaluating public-sector organizational networks. *Public Administration Review*, 61(4), 414-23.
- Provan, K.G., Fish, A., & Sydow, J. (2007). Interorganizational networks at the network level: A review of the empirical literature on whole networks. *Journal of Management*, 33, 479-516.
- Puterman, M.L. (1994). *Markov decision processes: Discrete stochastic dynamic programming*. New York: John Wiley & Sons.
- Rosen, R. (1999). *Essays on life itself*. New York: Columbia University Press.
- Rosen, R. (2005). *Life itself: A comprehensive inquiry into the nature, origin, and fabrication of life*. New York: Columbia University Press.
- Rumsfeld, D. (2003, June 24). *Joint capabilities integration development system* (CJCSI 3170.01C).

- Thompson, J. (1967). *Organizations in action: Social science bases of administrative theory*. New York: McGraw-Hill.
- US DoD Joint Forces Command. (2009). *Capstone concept for joint operations*. Retrieved February 1, 2009, from http://www.jfcom.mil/newslink/storyarchive/2009/CCJO_2009.pdf
- Weber, E.P., & Khademian, A.M. (2008). Wicked problems, knowledge challenges, and collaborative capacity builders in network settings. *Public Administration Review*, 68(2), 334-349.



A Technique for Evaluating Complex System of Systems Designs

Stephen Blanchette Jr.—Stephen Blanchette, Jr. is the Deputy Chief Engineer for Army programs at the Carnegie Mellon Software Engineering Institute in Pittsburgh, Pennsylvania, where he specializes in acquisition improvement initiatives. He has over 23 years' experience in the defense industry as a software engineer and manager, including prior positions with United Defense, Stanford Telecommunications, and McDonnell Douglas. He is an associate fellow of the American Institute of Aeronautics and Astronautics and a senior member of the Institute of Electrical and Electronics Engineers. Mr. Blanchette earned a BS in Computer Science from Embry-Riddle Aeronautical University and an MA in Diplomacy from Norwich University.

Mr. Stephen Blanchette, Jr.
Acquisition Support Program
Software Engineering Institute
4500 Fifth Avenue
Pittsburgh, PA 15213 USA
412-268-6275 (voice)
412-268-5758 (fax)
sblanche@sei.cmu.edu

Steven Crosson—Steven Crosson is the Associate Director (Acting) for Software in the Army's Program Executive Office Integration. He has been with the Army for 6 years and the federal government for 7 years. He has held several project lead and chief engineer positions including the Combat Net Radio and United States Message Text Format software testing projects for the Communications-Electronics Life Cycle Management Command Software Engineering Center as well as being the chief software engineer for PM Future Combat Systems. Mr. Crosson earned a BS in Computer Engineering from the University of Delaware, and an MS in Software Engineering from Monmouth University.

Mr. Steven Crosson
United States Army – Program Executive Office Integration
6501 E. 11 Mile Road
SFAE-FCS-E/MS: 515
Warren, MI 48397 USA
steven.crosson@us.army.mil

Abstract

Complexity is the hallmark of most modern military systems. The desire to have legacy systems interoperate with new systems, and especially the mounting interest in developing “systems of systems” (SoS) solutions, drive ever-more complexity into weapon systems. Complexity is further compounded by the increasing reliance on software to enable these systems. Existing forms of schedule- or event-driven reviews are inadequate to address the needs of software development in a complex SoS environment. What is needed is a true, evidence-driven, SoS-level evaluation capable of providing an overall assessment of, and insight into, the software development effort in that context. The Lifecycle Architecture anchor point is a technique used, typically, to evaluate the designs of single systems. This paper examines how, with some adaptation, the precepts of the Lifecycle Architecture anchor point can be scaled and applied to the system of systems domain.

Keywords: System of Systems, SoS, LCA, complexity



Acknowledgements

The authors gratefully acknowledge the contributions of Dr. Barry Boehm of the University of Southern California, who was instrumental in developing the technique described in this paper.

Introduction

Modern military systems are complex undertakings. The desire to have legacy systems interoperate with new systems, and especially the mounting interest in developing “systems of systems” (SoS) solutions, drive ever-more complexity. The problem is exacerbated by the reliance upon software to accomplish much of the underlying functionality; indeed, software often is the integrating element of an SoS. This complexity challenges the ability of engineers, managers, and users to achieve a solid understanding of an SoS during its development. Frequently, the SoS stakeholders (apart from the software developers themselves) have a very limited understanding of software or its contribution to the overall SoS. Yet, understanding an SoS during development is crucial to managing the development effort efficiently and to delivering quality products to the warfighter within schedule and budget.

Traditionally, schedule- or event-driven reviews have been a crucial element of most major development projects. From a software perspective, such reviews tend to focus on different aspects of development: producibility reviews focus on the ability to produce the software within available resources; capability reviews focus on the services being provided by the software; integration and test reviews focus on the readiness of the software to enter or transition between those phases of the development life cycle; schedule reviews focus on the development effort’s adherence to planned timelines; and so on. Although these different types of reviews provide valuable insights into the software development project at various stages of the lifecycle, the sum of these reviews is not sufficient to address the needs of software development in a complex SoS environment.

Most overall-systems reviews concentrate mainly on the functional definitions of system artifacts, and often treat evidence that the artifacts described will meet the system’s key performance parameter requirements comparatively lightly. Further, in order to support the needs of manufacturing planning and long-lead purchasing, such reviews tend to occur at points in the system development lifecycle ahead of the majority of the software development work. The resulting gap between maturity of SoS constituent systems and maturity of SoS software leads to uncertainty about the sufficiency of the overall SoS solution at a time when commitments must be made. Thus, traditional review approaches simply are inadequate when used as a means for understanding and evaluating software at an SoS level. How can stakeholders have confidence in an SoS that relies heavily upon software, when production commitments must be made well before completion of that software?

Prior to its 2009 cancellation, the US Army’s Future Combat Systems program faced just such a dilemma. Needed was a true SoS-level evaluation capable of providing an overall assessment of, and insight into, the software development effort in that context—one that could provide a review/assessment of how the developed software capability enabled the program’s required operational capability. As a solution, the program hypothesized that something like a Lifecycle Architecture anchor point review (often referred to simply as LCA), conducted at the SoS level, could answer the question. Originally a software development notion in the Rational Unified Process (RUP), the LCA marks the conclusion of



the elaboration phase of software development, when the requirements baseline is set and the architecture is complete (Kroll & Kruchten, 2003; Boehm, 1996). As extended to systems engineering in works by Richard Pew and Anne Mavor (2007) and Barry Boehm and Jo Ann Lane (2007), the goal of anchoring events such as the LCA is to assess program risk by examining evidence provided by the developers that a system developed to the architecture can satisfy the requirements. The LCA, in particular, strives to ensure that a system, as architected, can be constructed while also meeting planned cost and schedule targets.

This paper reports on the extension of the Lifecycle Architecture anchor point concept to the SoS level and describes its application to the software and computing elements of the former Future Combat Systems program.

The Lifecycle Architecture Anchor Point

Nominally, the LCA anchor point occurs at a stage in the development lifecycle where stakeholders evaluate the work that has been completed through the elaboration phase (the phase in which requirements and architecture models are largely complete) and assess the risks of moving forward into the construction and transition phases (the phases in which design, implementation, and testing/validation are performed). The name LCA derives from the nature of the review: it is a look-ahead through the lifecycle, conducted at a point where the architecture is sufficiently mature to allow reasoning about the relative risks of continuing to the construction phase of development. It is a foundational event, one where project stakeholders come together and agree to move forward, hence the term “anchor point”. The LCA differs from traditional milestone reviews such as preliminary design reviews (PDRs) and critical design reviews (CDRs), which tend to focus superficially on voluminous system description data, in that the LCA is a risk-based assessment focused on the feasibility of proceeding with work.

Central to the LCA is the notion of feasibility rationale, which documents evidence, provided by the developers, that the proposed architecture can be implemented to satisfy its requirements within the defined schedule and project budget. To justify the confidence of all stakeholders in moving forward into the latter phases of the development life cycle, the evidence must be both objective and internally consistent. The LCA is not just a technical assessment, but also a programmatic one. Successful completion of the LCA anchor point represents a commitment by all stakeholders to proceed with the program, based on objective evidence that the risks of doing so have been identified and sufficiently mitigated to provide a reasonable chance of project success. In contrast, insufficient or highly subjective evidence will not engender the confidence among stakeholders needed for them to commit to further development. Importantly, risks are addressed at a time when they are handled more easily and inexpensively than if they are allowed to propagate to later development stages.

The LCA feasibility rationale is relatively straightforward in the context of a traditional system development project. For a complex system of systems program, the feasibility package becomes less clear. Simply rolling up results from the LCA anchor points of constituent systems, for example, is not a sufficient approach because important inter-system and SoS implications might easily be missed. The feasibility of executing each individual software package/system does not necessarily correlate to the feasibility of executing the entire SoS.



Extending the LCA Process For The FCS SoS

Prior to its cancellation, the Future Combat Systems (FCS) program extended the notion of a Lifecycle Architecture anchor point to the SoS level as a means of evaluating on-going software development activities across the lifecycle in terms of risk with respect to meeting the program's Key Performance Parameters (KPPs).

The first step was to set up a team of experts from the Software Engineering Institute, the University of Southern California, and the Fraunhofer Center for Experimental Software Engineering at the University of Maryland (hereinafter referred to as the *SoS LCA Team*), with representatives from the Army providing guidance and representatives from the program's Lead System Integrator (LSI) facilitating access to program artifacts and personnel. Team members brought a range of programmatic and technical expertise to the effort, including a deep understanding of the usual LCA process for software. They also brought the degree of independence necessary to assure both the LSI and the Army of an unbiased result.

The SoS LCA Concept

In defining the SoS LCA review process, the SoS LCA Team followed a few key guidelines set by program management:

- Answer the following questions:
 - Can what is planned to be built really be built?
 - Will the various pieces add up to the desired whole?
 - Are the risks clearly identified and mitigated?
- Base answers on evidence rather than unsupported assertions.
- Discover issues early so that they can be fixed sooner rather than later.
- Build the confidence needed in the software/computing system development plans for a successful Milestone C decision for the overall program.

Rather than finding fault and assigning blame, the goal of the SoS LCA was to surface problems as early as possible so that they could be fixed at minimum cost. Equally important for FCS was the building of confidence in the software and computing system development plans, since an entire build of software would not yet have been developed at the time of the program's planned Milestone C review.

The FCS program had been executing traditional LCA reviews at the supplier and first-tier integration levels; however, simply rolling-up the results of these lower level reviews would have invited missing critical subtleties in the cross-system relationships that are the essence of an SoS. Also, the lower level LCA reviews focused narrowly on specific builds, often at different points in time. The SoS-level LCA was required to assess feasibility across future builds. Further, while other FCS software reviews focused on the plans for the immediately upcoming phase of a given build, the SoS LCA had to consider existing data and results from prior builds as well. The SoS LCA Team had to construct an SoS LCA evaluation process that had as its foundation the best-available facts and data as the basis for projecting forward.

As shown in Figure 1, the FCS SoS LCA process extrapolated, from what was known through what was expected and planned, the likelihood of reaching the desired endpoint within a tolerable level of risk. In the case of FCS, the desired endpoint was an



acceptable implementation of software Build 3 Final (B3F) to support a successful Milestone C decision (consequently, the SoS LCA results were a key feeder into the program's SoS PDR).

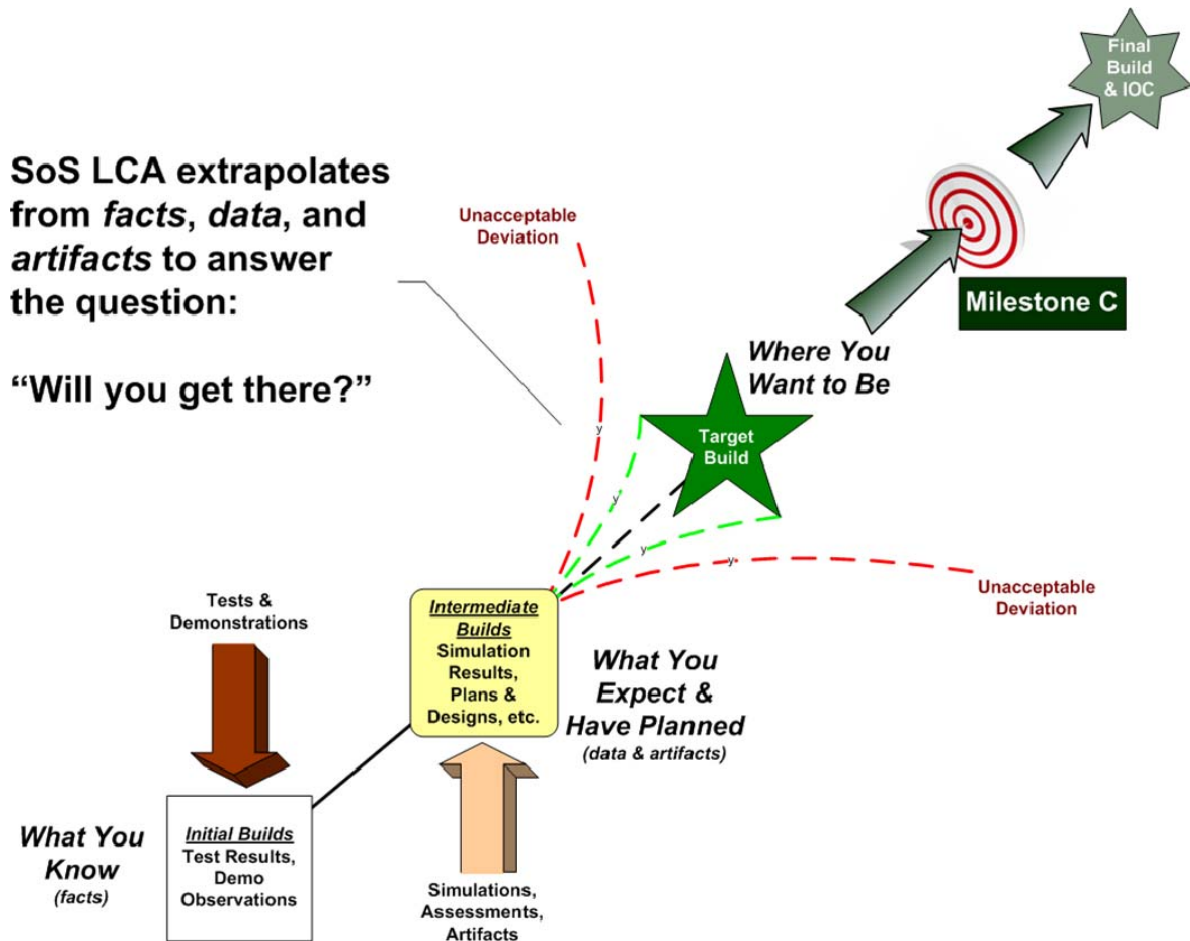


Figure 1. The SoS LCA Assesses the Likelihood of Achieving an Acceptable Outcome

Without waiting for tests on the final implementation, the SoS LCA sought to determine if FCS software could be built and if it would satisfy program needs. The basis of the extrapolation was objective evidence: *facts* and *data* in addition to other artifacts. Facts took the form of test results and demonstration outcomes for software that already had been built. Data consisted of simulation results that predicted operational performance of planned designs as well as the results of various technical assessments. The remaining artifacts of interest included the various plans, designs, schedules, and so on, that formed the basis of the work yet to be done. Engineering work products were evaluated for completeness and adequacy, as well as for consistency with other work products and plans. The SoS LCA focused on test data and results (as well as software producibility analysis) to evaluate the current capability of FCS software and to project the ability to meet program needs. The goal was to determine if development was on an acceptable trajectory or if deviations of sufficient significance required adjustments to development planning or execution.

Scope

On a very large SoS program such as FCS, there is no practical way to review all of the data one might wish to analyze in order to draw conclusions. Such a detailed effort might require thousands of labor hours spread over many months. While the cost alone would be a significant deterrent, the time factor also would be of concern. The longer it takes to complete the examination of all the evidence, the higher the risk that analyses performed early in the process will be invalidated by ongoing development work (due to prior problems or unknowns being resolved or new issues being introduced).

On FCS, program management had decided to limit the scope of the SoS LCA to the software and computer-processing elements of the program, but even that limitation left the range of investigation too broad. It quickly became apparent to the SoS LCA Team that executing a traditional LCA at the SoS level, even one constrained to the software and computing system elements of the program, would not be feasible. Merely coordinating an effort of that magnitude would become a project unto itself. Such an LCA would be too large, too complex, and too resource-intensive to be managed successfully. Instead, the team had to determine what a modest-sized team with limited resources could reasonably accomplish. The result was a focus on high-payoff areas for the program. These focus areas were not necessarily risk areas, but rather crucial areas to successful development of the FCS software:

1. ability to meet schedules within and across increments
7. ability to meet budgets within and across increments
8. ability to integrate across Integrated Product Teams (IPTs) and suppliers (including adequacy of requirements, architectural consistency, and so on)
9. ability to achieve/maintain interoperability among independently evolving systems (including Current Force platforms)
10. ability to coordinate multiple baselines for the core program, spin outs, and experiments
11. ability to manage co-dependencies between models, simulations, and actual systems
12. feasibility of meeting required performance levels in areas such as safety and security
13. maturity of technology considering scalability requirements
14. supportability of software
15. adequacy of commercial-off-the-shelf (COTS) evaluations and mitigations

A review of the results from lower level LCAs showed that these focus areas were commonly cited as problems, thus validating the list. With the scope bounded, the SoS LCA Team next worked on developing a process model for the SoS LCA.

An SoS LCA Process Model—First Cut

Figure 2 shows the initial SoS LCA process model. As envisioned, several analysis teams, consisting of the experts from the SoS LCA Team as well as representatives from the government, LSI, and the program's suppliers (referred to as One Team Partners, or OTPs), were to be formed. Each team would tackle a single focus area, examining the relevant facts, data, evidence, and plans and developing conclusions based on those



examinations. The results from each analysis team would be collected, summarized, and reported on by the SoS LCA Team, with a final report that detailed findings and recommendations serving as input to the program's SoS PDR. The SoS LCA final report, combined with results from the SoS PDR, would provide the basis for management decision making regarding any corrective actions.

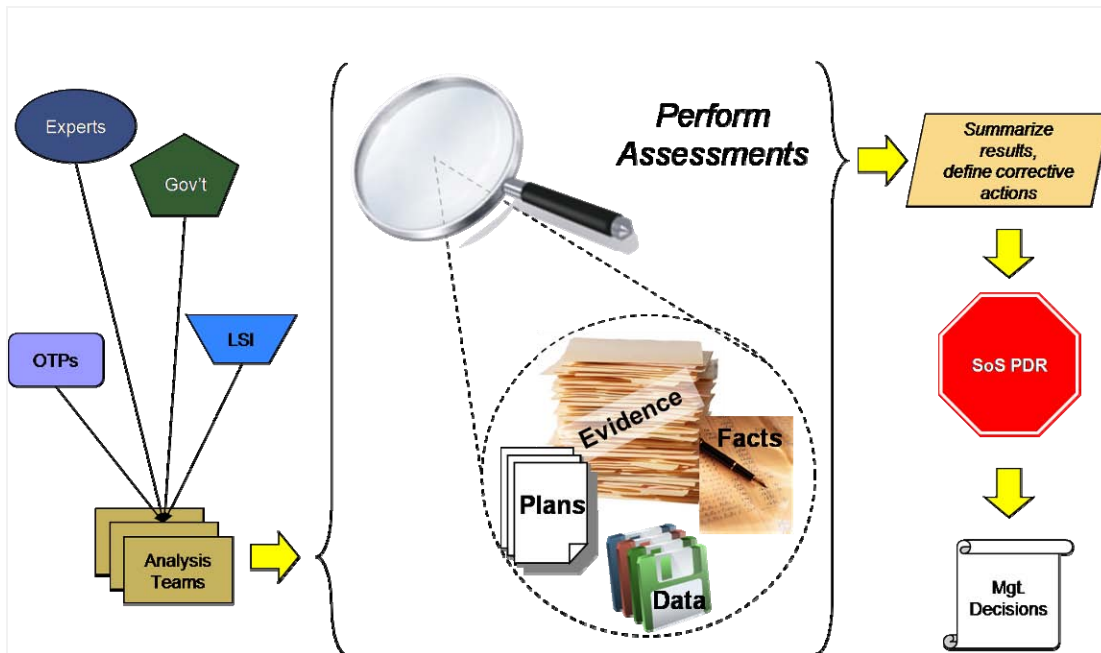


Figure 2. The Initial SoS LCA Process Model

Although the process model seemed to be straightforward, team members' experiences indicated that piloting the SoS LCA process on a small scale would be a wise first step. A pilot would allow the team to figure out the details of executing the process and discover any invalid assumptions and process inefficiencies.

The SoS LCA Pilot

Key considerations in planning for the SoS LCA pilot effort were its scoping and timing. It was important to limit the scope of the pilot to be executable relatively quickly, while at the same keep it broad enough to be useful as a test run for the larger effort. Equally challenging was scheduling the pilot. On a program with so many separate but interacting aspects, it seemed as if there was always a conflicting event to draw time, attention, and resources away from the SoS LCA pilot. As the pilot was not part of the original program plan, the need to minimize disruption to ongoing program activities served as yet another complication.

As it happened, the lead of the C4ISR¹⁹ team volunteered his area of responsibility to participate in the SoS LCA pilot because he saw value in having some independent review of the work being performed by his team. The C4ISR area, however, was quite sizeable with many interdependent elements, which was good from the perspective of piloting the process

¹⁹ C4ISR stands for command, control, communications, computers, intelligence, surveillance, and reconnaissance.

but challenging in terms of completing the pilot in a short timeframe. The SoS LCA Team negotiated with the C4ISR team lead to adjust the pilot scope to be achievable yet still worthwhile. The negotiated scope included a subset of the 10 SoS LCA focus areas, coupled with the C4ISR lead's principle areas of concern. The timing of the pilot also aligned well with a previously planned review cycle for the C4ISR team.

After first socializing the notion of the pilot among the C4ISR team members so that expectations on both sides were clear, the SoS LCA Team next set about gathering data. Team members divided the focus areas among themselves based on their respective backgrounds and expertise. The Team spent about three weeks gathering evidence from existing C4ISR plans and technical documents; no new documents were to be created for the purposes of the pilot.

Almost immediately, the pilot began serving its purpose. One early discovery was that SoS LCA Team members were not in full agreement about the goals or method of conduct of the pilot. Thus, the pilot was an invaluable experience both from the perspective of checking and refining the process and from the point of view of developing a common mission understanding within the SoS LCA Team.

Another revelation was the considerable amount of misunderstanding between the SoS LCA Team, the C4ISR team, and its various performer organizations (despite efforts to set clear expectations at the outset). Documentation and evidence furnished by the C4ISR team generated more questions than answers, requiring numerous follow-up teleconferences and requests for additional documents. The documents provided often had timestamps differing by several months and frequently made what appeared to be incompatible assumptions. Trying to resolve these inconsistencies without interrupting the performers' work gave the SoS LCA Team an appreciation of the difficulty to be expected when performing a complete assessment.

The SoS LCA Team found little documentation to support an independent conclusion that the C4ISR plans and approach were feasible. This outcome was not because the C4ISR team had not done its job; on the contrary, there was ample evidence of significant engineering effort. However, there was no contractual obligation for suppliers to demonstrate the feasibility of their work, per se, outside of customary test and validation methods, which placed a heavier burden on the analysis team in trying to understand how completed work related to future work. Since the crux of the LCA is reliance upon evidence rather than verbal accounts and assurances, the SoS LCA Team realized that the full assessment would require extra time for data gathering and a deeper probe of the available evidence in order to ensure the veracity of findings.

Despite the noted difficulties, the pilot generally was a success in that it helped clarify the SoS LCA Team's thinking about conducting the full-scale SoS LCA. It demonstrated that program personnel would be too busy to participate significantly in gathering and examining feasibility evidence, which would necessitate a larger effort on the part of the analysis team. It also demonstrated the need to set expectations among developers and management continually to ensure their ongoing support of the process. The pilot also highlighted the need to review emerging findings with stakeholders to ensure accuracy. Meanwhile, the C4ISR Team also benefitted from the pilot by being able to raise issues that had been of concern to them for management attention, with the added weight of independent analysis to support them.



Some Process Refinements

As a result of the SoS LCA pilot and other program activities, the SoS LCA Team made some adjustments both to the focus areas and to the process model.

First, in consultation with program management, the SoS LCA Team modified the focus areas. The initial list proved to be too vague and overly oriented toward programmatic (e.g., non-technical) considerations. The modified list represented the essential FCS capabilities, the "non-disposable" elements of the program. These elements (see Table 1) were ones that had to work in order to ensure that the program could produce a viable SoS capability for the Warfighter.

Table 1. Final SoS LCA Focus Areas

Focus Area	Description
Fusion and Data Distribution	Includes: <i>Distributed Information Management</i> —effectiveness of data communications among platforms and systems over a network, analysis of data reliability, latency, etc. <i>Distributed Fusion Management</i> —effectiveness of utilizing several sources to gather and distribute fusion data, including network usage, data reliability, latency, etc. <i>Rapid Battlespace Deconfliction</i> —effectiveness of providing, recognizing, and utilizing battlespace deconfliction data passed among Brigade Combat Team (BCT) platforms
Quality of Service (QoS)	Analysis of message prioritization effectiveness, delivery rates, methods of prioritization, etc.
Information Assurance (IA) & Security	Review of the status of efforts to develop and test FCS IA components, and a determination of the attainment of necessary functional capabilities
Software Performance	Analysis of the projected ability of fielded FCS software (e.g., algorithms) to enable planned capabilities in a comprehensive, end-to-end operational environment
Display Management and Task Automation/Task Integration Networks (TINs)	Analysis of the performance, usability, and automation capabilities demonstrated in Warfighter Machine Interface (WMI) displays
Key Battle Command Systems	Includes: <i>Network Management System (NMS)</i> —review of current status of development of the Network Management System (NMS) system, as well as analysis of the current state of general network management issues <i>Integrated Computer System (ICS)</i> —assessment of the development and issues related to the various versions of the ICS system. Determination of the ICS' ability to meet FCS design needs <i>Centralized Controller (CC)</i> —analysis of the current state of the CC and its ability to meet necessary functional capabilities

The results of investigations into these focus areas formed the detailed technical basis for the final report to management.



Second, the realities of personnel availability necessitated slight adjustments to the process model. Figure 3 shows the changes, which were encapsulated entirely within analysis team formation (i.e., inside the red oval denoted by the gray arrow on the left-hand side of the diagram).

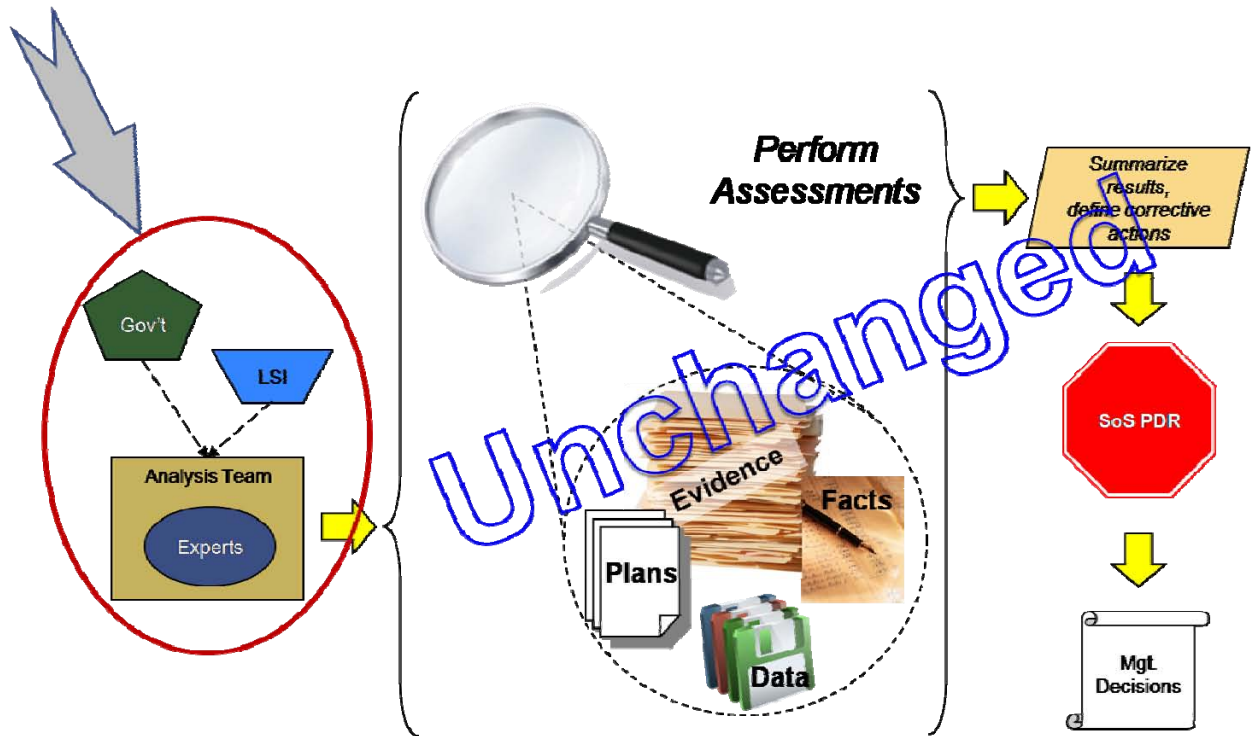


Figure 3. SoS LCA Process Model Showing Refinements

As shown in Figure 3, instead of several analysis teams, there was only one, which was composed entirely of technical experts from the SoS LCA Team, although Army representatives provided guidance and an LSI representative facilitated access to documentation. Suppliers were unable to participate in the analysis process at all, although they were called upon to provide data and, occasionally, clarification of the data. The SoS LCA Team divided the focus areas among themselves; in each area, typically one or two team members had responsibility for identifying the types of artifacts needed for evaluation, reviewing them, and rendering recommendations based on review findings. The entire analysis team reviewed results as they emerged and briefed status to program management at frequent intervals.

Capstone Analyses: Tying It All Together

While detailed technical analyses are the foundation of the SoS LCA, they are inadequate for communicating status and risk to stakeholders who may have limited expertise in those technical areas. In particular, the highest levels of program management need to be able to relate the SoS LCA findings to overall program outcomes in order to make informed decisions about risk mitigation. Further, the thread to tie the focus areas together in a meaningful way was missing. Needed was a means of looking at the results in a holistic manner and interpreting them in way that would be relevant to management

decision-making. The solution involved two capstone analysis efforts, one from a technical perspective and the other from a cost and schedule perspective. The end-state design analysis summed up the findings of the technical focus areas in a way that made the technical findings relevant for management decision-making. The software producibility analysis looked at the feasibility of developing the remaining software within allocated schedules and budgets based on a cost/schedule analysis of the completed builds cross-compared to the relative risks that emerged from the end-state design analysis. Taken together, these two capstone efforts summarized the overall SoS LCA findings in terms of program performance, not simply software performance, providing management with the “so-what” in a report that was otherwise highly technical.

End-State Design Analysis

The capstone technical analysis was an examination of the overall “end-state” design – the software design as it was expected to exist at the completion of the development program. The analysis sought to determine if the end-state software design would truly meet operational needs of the program. The analysis was an ambitious undertaking; there had never before been an attempt to understand the FCS SoS software design to such a depth. Two problems popped up immediately: 1) how to characterize operational needs in terms of software and 2) how to tie the analysis to those needs.

Rather than recast the program’s operational needs in software terms, a more logical approach to the first problem was to analyze the software contributions to the definitive characterization of the program’s operational needs—the SoS *key performance parameters* (KPPs). That realization made the analysis approach obvious. The SoS LCA Team applied *assurance cases* to analyze the SoS end-state software design with respect to the KPPs.

Assurance Cases

An assurance case is nothing more than a structured set of arguments, supported by a body of evidence, that justifies belief that a given claim is true. To make one’s “case,” one argues that certain evidence supports (or does not support) a given claim. One constructs an assurance case by starting with an overarching claim and then iteratively decomposing it into constituent claims, which, at the lowest level, are supported directly by evidence. Figure 4 depicts the concept.



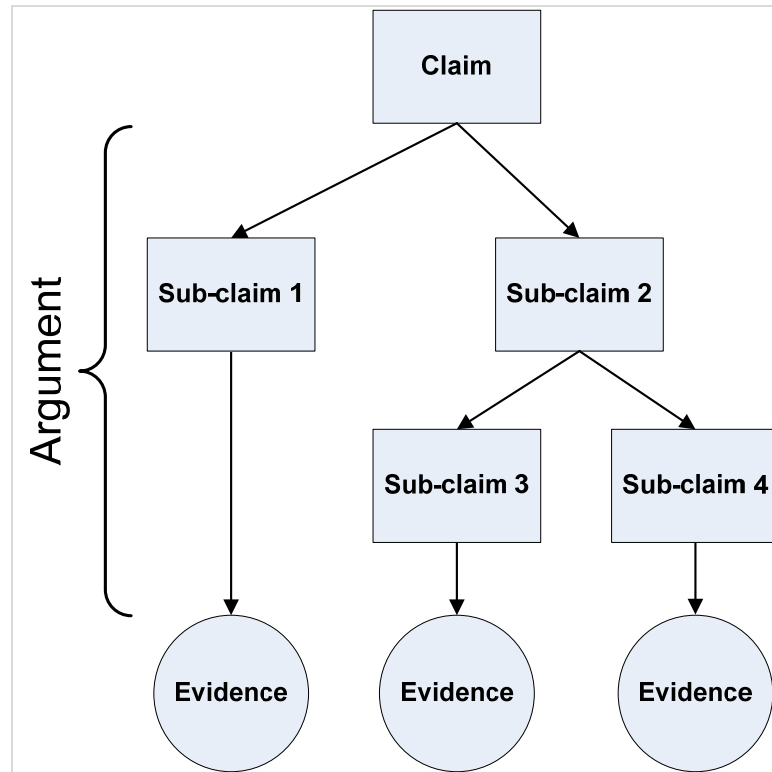


Figure 4. The Basic Structure of an Assurance Case

The logic of an assurance case is straightforward. If all the evidence shown in Figure 4 is sufficient, then belief in sub-claims 1, 3, and 4 is justified. Belief in sub-claims 3 and 4 justifies belief in sub-claim 2, which, combined with belief in sub-claim 1, justifies belief in the top-level overall claim.²⁰

Applying Assurance Cases—An Example

For the FCS SoS LCA, the end-state design analysis approach was to use each of the program's KPPs as a main claim and demonstrate how well the SoS software and computing system designs supported them. (In one case, software and computing systems played no role in satisfying a KPP; developing a partial assurance case confirmed that suspicion.) The findings from the focus area analyses served as primary sources of evidence, augmented with additional analyses and artifacts as necessary.

Figure 5 presents an oversimplified example (note that it is not an actual FCS analysis).

²⁰ A more complete discussion on the use of assurance cases for analyzing SoS software designs can be found in "Assurance Cases for Design Analysis of Complex System of Systems Software" (Blanchette, 2009).

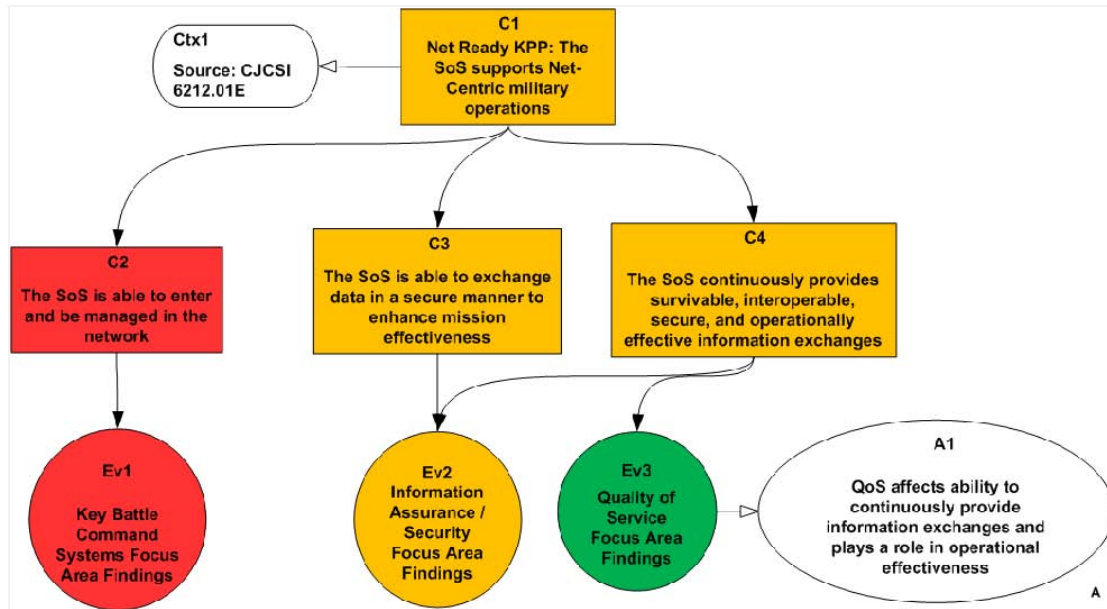


Figure 5. Oversimplified Example of an Assurance Case Using Focus Area Findings as Evidence

The example shown is based on the Net Ready KPP, a common KPP among DoD systems that exchange information. As shown by the context bubble, Ctx1, the KPP definition is taken from the Chairman of the Joint Chiefs of Staff Instruction (CJCSI) 6212.01E (Chairman of the Joint Chiefs of Staff, 2008) rather than on a specific FCS program requirement. The figure shows the main claim, labeled C1, is supported by three sub-claims, labeled C2, C3, and C4. For purposes of illustration, each sub-claim is shown to be supported directly by findings from the SoS LCA focus areas (Ev1, Ev2, and Ev3).²¹ An actual analysis would be far more extensive, with many sub-claims and possibly multiple sources of evidence supporting the claims at the lowest level.

For purposes of communicating findings to management, it is useful to speak in terms of risk (i.e., risk of not meeting a KPP). A common management practice is to associate a color code with different risk levels, such as red to indicate a high level of risk, yellow to indicate a moderate level of risk, and green to indicate a low level of risk. The relative risk levels are then rolled up according to predetermined rules to establish a level of confidence about the supported claims. A sample rule set for this example is shown in Table 2.

²¹ In this sample case, assumption A1 helps to justify the use of the QoS focus area findings in support of claim C4 since the connection might not have been obvious.

Table 2. Sample Rules for Rolling Up Risk

Green (low risk)	All lower level claims and supporting evidence are green.
Yellow (moderate risk)	Some lower level claims and supporting evidence are a combination of yellow and red.
Red (high risk)	All, or an overwhelming majority of, lower level claims and supporting evidence are red.

To further illustrate the assurance case concept, the authors have arbitrarily assigned one of each color-coded risk level to the evidence bubbles. Applying the rule set to the example, the findings for key battle command systems indicate a high risk within the focus area (Ev1), so the claim supported by that evidence is also rated as high risk (or, more correctly, low confidence that the claim is true). Applying the rule set to the remaining elements of the example yields the results shown. Since none of the sub-claims in the example is rated any better than moderate risk (indeed, C2 is high risk), the main claim C1 can be no better than moderate risk and, therefore, there is only moderate confidence in satisfying the Net Ready KPP in this example.

A full analysis, of course, would identify the specific issues that led to the risk assignments as well as corresponding recommendations for addressing them. Given such information, a program manager could decide if the moderate risk of not meeting the KPP was acceptable or, if not acceptable, decide on one or more courses of action for reducing the risk/increasing confidence. Although this example is deliberately negative, it should be clear that strong evidence in support of all claims would lead to a high level of confidence that a given KPP was well supported; this situation is, of course, the ideal.

For FCS, the use of assurance cases provided a way to report the technical findings from the SoS LCA to program management in a way that related directly back to program goals without requiring a detailed understanding of software and computing systems. Each KPP had its own assurance case, and the findings from the technical focus areas were used as evidence to support one or more of those assurance cases. This approach tied both issues and strengths in the software to the overall program goals, providing a view of program risk that was vertically integrated within each KPP and horizontally integrated across the technical focus areas.

Producibility Analysis²²

The second capstone, producibility analysis, concerned the programmatic aspects of the FCS SoS software development effort. Apart from technical feasibility, the producibility analysis sought to demonstrate the feasibility of developing the SoS software and computing systems within cost and schedule targets.

Estimating software producibility presented some special challenges. Unlike hardware, which is typically built to completion before testing, the FCS software was being implemented in a series of increments or builds, which is a fairly common approach for large development programs. In such cases, the software producibility costs for later development increments tend to increase due to breakage in previous increments as well as increased integration and test costs.

²² Dr. Barry Boehm of the University of Southern California is the principal author of this section.

Breakage in previous increments has several sources. One source is errors discovered in the earlier increments that must be fixed within the budgets and schedules of later increments, which then increases development workload, while the integration and test burden for later increments is already larger due to the software product becoming more complete. Other sources of breakage include revisions of earlier-increment software to support later-increment needs or changes, and adaptation of the software to changes in COTS or external-software interfaces.

The magnitude of these Incremental Development Productivity Decline (IDPD)²³ effects varies due to several factors, such as the degree of coupling to other parts of the software or the rate of requirements or interface changes. Thus, using constant-productivity estimates for future software increment producibility projections would lead to severe underestimation.

Further, FCS was in a unique position in that there were no predecessor projects of equivalent scope and scale from which to predict accurately its software producibility rates. The estimation parameters and knowledge bases of current software estimation tools are generally good for stable, standalone, single-increment development of the kind of software being developed at the lower system levels by the FCS suppliers. However, such tools generally fail to account for the degrees of program and software dynamism, incrementality, coordination complexity, and system of systems integration that were faced by FCS. These phenomena tend to decrease software productivity relative to the cost model estimates, but it is difficult to estimate by how much.

To calibrate software estimates involves measurement of producibility data from early increment software deliveries. For FCS, the best available relevant data was from the System of Systems Common Operating Environment (SOSCOE) software because it underpinned most other mission software packages and thus had several early releases; consequently, there were four increments upon which to base the magnitude and evolution of its IDPD factor. SOSCOE was not fully representative of the mission software, however, leaving uncertainties about its use to predict the IDPD factors of the mission software increments. Similarly, mission software from other programs was not representative of the unique system of systems aspects of FCS. Thus, data from both sources (SOSCOE and other programs) were used to perform a sensitivity analysis of the amount of software that could be produced within the available budget and schedule.

Research Results

Existing software reviews had specific artifacts and criteria to examine, but their scope was necessarily limited. By defining appropriate and relevant criteria at the SoS level, the SoS LCA provided the opportunity to independently assess broader areas that were critical to FCS success. Anchoring the evaluation results with facts and data from early software builds provided a powerful and objective basis for judging the realism of plans for future builds, both from technical and productivity perspectives. Findings were refutable if engineers were able to supply evaluators with additional evidence that had not been

²³ IDPD refers to the phenomenon in which software developer productivity tends to decrease during the later increments of an incremental development project due to a larger code base that must be maintained and integrated in addition to producing new code for the current increment (Boehm, 2009).



considered;²⁴ indeed, there was one instance in which evaluation findings were revised based on evidence that had not been provided previously. Equally important, the evidence basis of the SoS LCA protected against pessimistic appraisals. In the end, program engineers were largely in agreement with the SoS LCA Team's findings.

The producibility and end-state design capstone analyses (which included roll-ups of findings from the technical focus areas) indicated several places where key changes needed to be made to the FCS SoS software development effort. What is more, regular reports of preliminary findings helped management make adjustments to the program or, in some cases, to the in-progress evaluations. Gaps in design, architecture, and requirements coverage that were unlikely to have been uncovered through other types of reviews were identified and prioritized for correction, providing a path to reach necessary capability levels in time to support a Milestone C decision. Equally important, the SoS LCA provided a mechanism to report technical, cost, and schedule risks, clearly tied to overall program goals, at an appropriate level of detail for senior program management, thereby enabling their understanding and supporting their decision-making processes.

The FCS SoS LCA experience was extremely valuable in providing insights into processes better fitted to the DoD's future systems of systems programs. For example, it has stimulated several improvements to the DoD version of the Incremental Commitment Model (ICM), initially defined by Pew and Mavor (2007) as a process model for DoD human-intensive SoS. The resulting upgrade expressed in its draft DoD version changes the name of the pre-development phase from "Architecting" to "Foundations" (including plans, budgets, schedules, and contract provisions as well as architecture), and changes the name of its phase end-gate from "Lifecycle Architecture" to "Development Commitment Review" (Boehm & Lane, 2008). The SoS LCA also stimulated extension of the ICM process to include guidance for planning, executing, and earned-value managing the development of feasibility evidence as a first-class deliverable. Complementary process improvement initiatives for systems of systems are underway at the SEI.

Some Lessons

FCS conducted the SoS LCA prior to its SoS PDR, which was a good fit for the program; it enabled software to have a relevant voice during an event that otherwise would have been dominated by system and SoS concerns. Had the program continued, the authors believe that a follow-up SoS LCA would have been advantageous prior to the planned CDR as well.²⁵ In fact, the Brigade Combat Team Modernization (BCTM) program, a follow-on to FCS, is exploring how best to utilize another SoS LCA. In general, the SoS LCA approach described here could be applied at either or both points in the development lifecycle as shown in Figure 6; the key determining factors are the maturity of the SoS architecture and designs, and the availability of early implementations to provide an analysis baseline.

²⁴ Note that such evidence had to have been in existence at the time of the evaluation, not created after the fact to defend against the findings.

²⁵ In theory, it should be possible to execute a similar type of analyses milestone (called an Initial Operational Capability anchor point in the RUP), again scaled to the SoS level, at the conclusion of the construction phase of development but still prior to initial operational testing, although the need for doing so and the impact of the outcomes are subjects for further research.

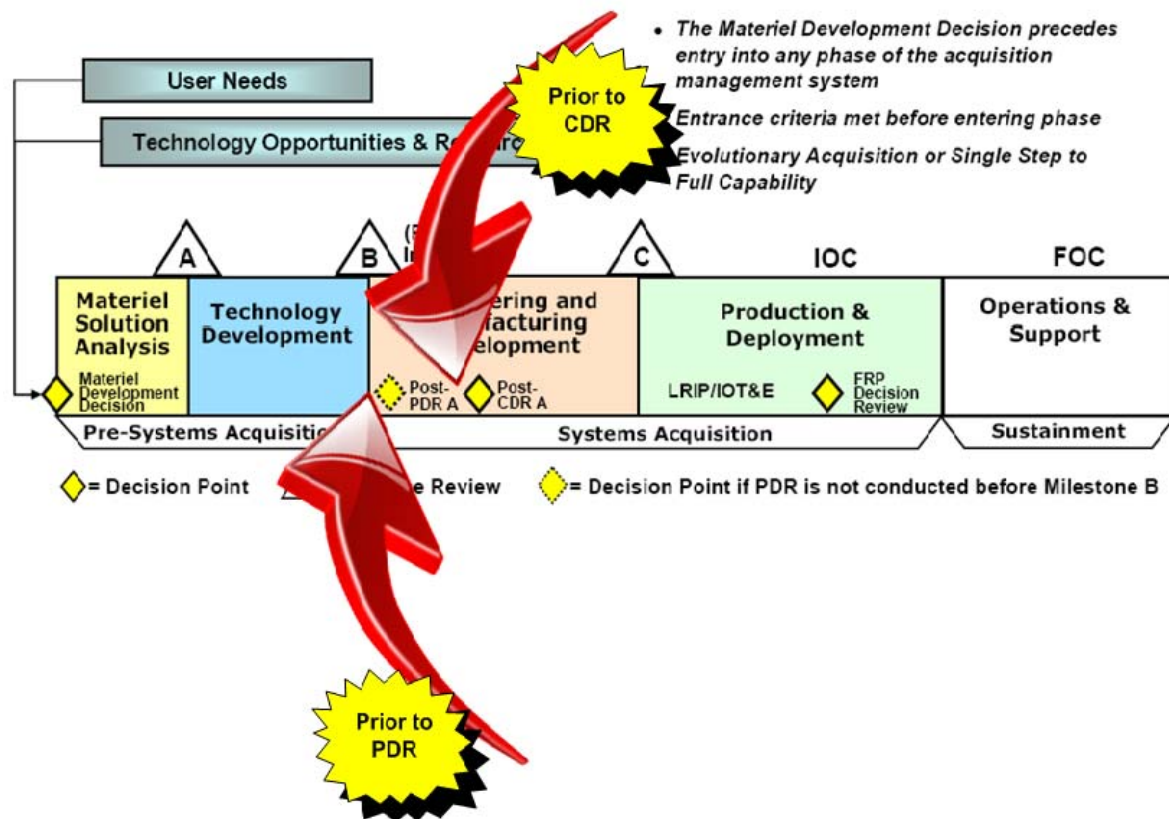


Figure 6. The SoS LCA Can Be Employed Prior to Traditional Milestone Reviews

Obtaining contractual commitment to have the feasibility evidence produced and the independent evaluation²⁶ performed is essential. Without such commitment, the SoS LCA Team's requests for data simply did not have sufficient weight when balanced against other program priorities. Once the SoS LCA became a contractual requirement, it was easier for program management at all levels to prioritize requests for information appropriately. In addition, management commitment is indispensable to achieving benefit from an SoS LCA. The true value of the SoS LCA is in being able to make informed program decisions based on facts and data. While the intent is always to conduct an SoS LCA in a way that facilitates decision-making, without true management support the effort (like many other program reviews) is at risk of becoming a "check the box" exercise, where the event is held in order to claim completion on a schedule without regard for the evaluation's effectiveness. Here also, contractual provisions and incentives for developers to plan for and produce feasibility evidence as a first-class deliverable are important elements.

²⁶ Technically, there is no requirement for independent parties to conduct an LCA evaluation. For the FCS program, using independent experts allowed the SoS LCA to be inserted into the development effort with minimal cost and disruption (such as for training staff). Had the SoS LCA been part of the program plans from the outset, independent experts might not have been needed as program staff could have been trained to produce and evaluate the necessary feasibility rationale.

To facilitate inclusion of an SoS LCA in a contract, standards for feasibility evidence should be developed. Such standards would make clear to all parties the kinds of evidence and levels of detail that would be acceptable, enabling the developers to produce it and evaluators to assess it. The FCS program had done this for its lower-level LCA reviews, making the standards a contractual obligation on suppliers. However, the relatively late idea to perform an SoS-level LCA left no time to develop appropriate standards at the SoS level. While this problem was overcome through deeper subject matter analysis, having agreed upon standards would have lessened the burden on the team performing the analyses. Exactly what those standards ought to be is not clear. Surely, the lessons from FCS could be applied to the follow-on BCTM program, but there likely is a generic set of standards that could be extrapolated to fit many different types of SoS.

Gathering artifacts took much longer than anyone had anticipated, during both the pilot and the formal SoS LCA. Part of the difficulty lay in communicating exactly which artifacts were needed and which were available. While the SoS LCA Team believed it communicated its needs clearly, differences in interpretation between the team and program personnel providing the artifacts caused confusion and slowed the process. Another difficulty was that, owing to the sheer size of the program, just finding the custodian who had access to needed information often took a long time.

Keeping management informed of progress and emerging results is essential, but doing so must be managed carefully. As findings became available, status updates to management became more frequent, which in turn led to a decrease in analysis due to the need to spend time to prepare and rehearse briefings and provide read-ahead material in advance of the actual update meetings. Management updates should be scheduled at set points during SoS LCA execution as part of the process.

“Socializing” the process among stakeholders and management in advance is crucial to setting expectations. Many people initially found the SoS LCA concept obscure. Indeed, the term “Lifecycle Architecture” caused a great deal of confusion among managers and project personnel when applied in an SoS context. The word *architecture*, in particular, sets incorrect expectations that the event will be nothing more than an architectural review, which is neither the intent of an LCA review (at any level) nor, as explained earlier, feasible at the SoS level due to the differing development stages among the constituent systems. For this reason, a name other than “LCA” (e.g., Development Commitment Review as used in the Incremental Commitment Model) should be considered while still maintaining the principles embodied in the SoS LCA concept.

Lastly, while the FCS program applied the SoS LCA to software and computing systems only, it should be possible to use the technique on a broadened technical scope encompassing hardware and system issues as well. As a practical matter, it is nearly impossible to ignore such issues even with a limited focus on software.

Conclusion

Extending the concept of an LCA review process to the SoS level was an effective means of evaluating the FCS SoS software and computing system designs. Both the depth and breadth of analysis of this software review far exceeded other software-specific reviews on the program. The broad, multi-build SoS view, in conjunction with the individual build reviews provided an excellent assessment of the state of the FCS software development effort and its potential for achieving program objectives. The technique provided insight into areas of the software development program that had never had an in-depth review.



Overall, the SoS LCA was a success both as a review technique and, more importantly, in providing a solid understanding of the functional baseline for FCS software leading into the SoS PDR. The analyses not only helped discover problem areas, but also recognized software packages that were meeting or exceeding expectations. This latter category was particularly important, as it provided guidelines for recommended paths forward for other software packages. Recommended changes to program processes and evaluations were proposed and being developed for inclusion in the follow-on BCTM program. Rather than mere action items to be tracked to answer specific questions, the SoS LCA provided more details and greater program benefit, including new direction for certain program areas.

The key element of the SoS LCA was the ability to report technical, cost, and schedule risks in a coordinated fashion relative to program goals and *at an appropriate level of detail* for senior program management, thereby facilitating their understanding and supporting their decision-making processes.

The success of the effort suggested possible follow-up activities had FCS continued, including performing one or more “delta” SoS LCA reviews to re-evaluate areas where the highest risks were found and inserting a similar type of effort before a program CDR to gain deeper insights into the development efforts at that stage of the program. While these specific actions were not tried on FCS, it seems reasonable to conclude that they would have been beneficial. This idea is consistent with the concept that shortfalls in feasibility evidence are uncertainties and risks, and should be covered by risk mitigation plans.

The FCS SoS LCA activity also has been a valuable learning experience for preparing and conducting future DoD SoS milestone reviews and acquisition processes.

References

- Blanchette, S., Jr. (2009). Assurance cases for design analysis of complex system of systems software. In *Proceedings of the AIAA Infotech@Aerospace Conference and AIAA Unmanned...Unlimited Conference*. Seattle: American Institute of Aeronautics and Astronautics.
- Boehm, B. (1996). Anchoring the software process. *IEEE Software*, 13(4), 73-82.
- Boehm, B., & Lane, J. A. (2007). Using the Incremental Commitment Model to integrate system acquisition, systems engineering, and software engineering. *CrossTalk*, 20(10), 4-9.
- Boehm, B., & Lane, J. A. (2008). *Guide for using the Incremental Commitment Model (ICM) for systems engineering of DoD projects, version 0.5* (USC-CSSE TR-2009-500). Center for Systems and Software Engineering, University of Southern California, Los Angeles, CA. Retrieved from http://csse.usc.edu/csse/TECHRPTS/by_author.html#Boehm
- Boehm, B. (2009). Future challenges for software data collection and analysis. In *Proceedings of the Predictor Models In Software Engineering (PROMISE) 2009*, <http://promisedata.org/pdf/2009/keynote.pdf>
- Chairman of the Joint Chiefs of Staff, (2008). *Interoperability and supportability of information technology and national security systems* (CJCSI 6212.01E). Washington, DC: Retrieved from http://www.dtic.mil/cjcs_directives/cdata/unlimit/6212_01.pdf
- Kroll, P., & Kruchten, P. (2003). *The Rational Unified Process made easy: A practitioner's guide to the RUP*. Boston: Addison-Wesley.
- Pew, R. W., & Mavor, A. S. (Eds.). (2007). *Human-system integration in the system development process*. Washington, DC: National Academies Press.



System Development and Risk Propagation in Systems-of-Systems

Muharrem Mane—Muharrem Mane is a postdoctoral researcher in the School of Aeronautics and Astronautics Engineering, Purdue University. He received his PhD from Purdue University in Aerospace Engineering in 2008. His current research interests are in risk analysis and propagation, resource allocation and design under uncertainty, and network modeling and analysis. He currently works in the System-of-Systems Laboratory led by Dr. DeLaurentis.

Muharrem Mane
Postdoctoral Researcher, School of Aeronautics and Astronautics
Purdue University
701 W. Stadium Avenue
West Lafayette, IN 47807
Phone: (765) 494 7958
Fax: (765) 494-0307
mane@purdue.edu

Daniel DeLaurentis—Daniel DeLaurentis is an Assistant Professor in the School of Aeronautics and Astronautics Engineering, Purdue University. He received his PhD from Georgia Institute of Technology in Aerospace Engineering in 1998. His current research interests are in mathematical modeling and object-oriented frameworks for the design of system-of systems, especially those for which air vehicles are a main element, and approaches for robust design, including robust control analogies and uncertainty modeling/management in multidisciplinary design.

Daniel DeLaurentis
Assistant Professor, School of Aeronautics and Astronautics
Purdue University
701 W. Stadium Avenue
West Lafayette, IN 47807
Phone: (765) 494 0694
Fax: (765) 494-0307
ddelaure@purdue.edu

Abstract

The emphasis of the Department of Defense on capability-based acquisition has led to the simultaneous development of systems that must eventually interact within a system-of-systems. Thus, system development and acquisition processes encounter interdependencies that generate complexity and risk. The authors' prior work has developed a Computational Exploratory Model to simulate the development processes of these complex networks of systems intended for a system-of-systems capability. The model's goal is to understand the impact of system-specific risk and system interdependencies on development time. The progress documented in this paper focuses on the quantification of risk propagation and the impact of network topologies on the propagation of disruptions. The improved model enables trade studies that differentiate the effectiveness of alternate configurations of constituent systems and that quantify the impact of varying levels of interdependencies on the timely completion of a project that aims to achieve a desired capability level.



Introduction

The purpose of capabilities-based acquisition, as described by Charles and Turner (2004), is to acquire a set of capabilities instead of acquiring a family of threat-based, service-specific systems. The Missile Defense Agency (MDA), for example, uses capability-based acquisition to evaluate the success of a program based on its ability to provide a new capability for a given cost, and not on its ability to meet specific performance requirements (Spacy, 2004). The Joint Mission Capability Package (JMCP) concept is another example that aims to create a joint interdependency between systems to combine capabilities in order to maximize reinforcing effects and minimize vulnerabilities (Durkac, 2005). The goal is a more efficient utilization of both human and machine-based assets and, in turn, improved combat power.

To accomplish the desired capability, systems are increasingly required to interoperate along several dimensions that characterizes them as systems-of-systems (SoS) (Maier, 1998). Systems-of-systems most often consist of multiple, heterogeneous, distributed systems that can (and do) operate independently but can also collaborate in networks to achieve a goal. Examples of systems-of-systems include: civil air transportation (DeLaurentis, Han & Kotegawa, 2008), battlefield ISR (Butler, 2001), missile defense (Francis, 2007), etc. According to Maier (1998), the distinctive traits of operational and managerial independence are the keys to making the collaboration work. The network structure behind the collaboration, however, can contribute both negatively and positively to the successful achievement of SoS capabilities and, even earlier, to the developmental success. Collaboration via interdependence may increase capability potentials, but it also contains concealed risk in the development and acquisition phases. Brown and Flowe (2005), for instance, have investigated the implications of the development of SoS to understand the drivers that influence cost, schedule, and performance of SoS efforts. Results of their study indicate that the major drivers—as indicated by subject-matter-experts—include systems standards and requirements, funding, knowledge, skills and ability, system interdependencies, conflict management, information access, and environmental demands.

Disruptions in the development of one system can have unforeseen consequences on the development of others if the network dependencies are not accounted. The goal of a single system's program manager is the mitigation of risk, leading to successful development of that specific system. While direct or immediate consequences of decisions are nearly always considered, the cascading second-and-third order effects that result from the complex interdependencies between constituent systems in an SoS are often not, which make success all the more difficult. It falls on acquisition managers and systems engineers (or systems-of-systems engineers) to understand and manage the successful development of a system, or family of systems, to produce the targeted capability in this challenging setting.

Evidence is abundant that system-of-systems-oriented endeavors have struggled to succeed amidst the development complexity. The Future Combat System is a latest example (Gilmore, 2006). Civil programs have not been spared either, e.g., Constellation Program (Committee on Systems, 2004) and NextGen (2009). Rouse (2001) summarizes the complexity of a system (or model of a system) as related to the intentions with which one addresses the systems, the characteristics of the representation that appropriately accounts for the system's boundaries, architecture, interconnections and information flows, and the multiple representations of a system.



The work presented in this paper specifically targets complexities stemming from system development risk, the interdependencies among systems, and the span-of-control of the systems or system-of-systems managers and engineers. *The objective of the research summarized in this paper is to quantify the impact of system-specific risk and system interdependency complexities using a computational exploratory modeling approach.* The work comprises new improvements to a computational exploratory model (CEM)—a discrete event simulation model—previously introduced in prior Acquisition Symposia (Mane & DeLaurentis, 2009) that aims to provide decision makers with insights into the development process by propagating development risk in the SoS network and capturing the impact that system risk, system interdependencies, and system characteristics have on the timely completion of a program. We also briefly introduce complementary work related to an analytical approach to treat the same complexities via computations on conditional probabilities that relate the transmission of risk in network dependent systems.

Computational Exploratory Model (CEM) Overview

The CEM is based on the 16 basic technical management and technical system-engineering processes outlined in the *Defense Acquisition Guidebook* (DoD, 2008a), often referred to as the 5000-series guide. However, an SoS environment changes the way these processes are applied. The *Systems Engineering Guide for System-of-Systems (SoS-SE)* (DoD, 2008b) addresses these considerations by modifying some of the 16 processes in accord with an SoS environment. The resulting processes and respective functions consist of translating inputs from relevant stakeholders into technical requirements, developing relationships between requirements, designing and building solutions to address requirements, integrating systems into a high-level system element, and performing various managing and control activities to ensure that requirements are effectively met, risks are mitigated, and capabilities achieved.

The CEM, centered on these revised processes, is a discrete event simulation of the development and acquisition process. This process creates a hierarchy of analysis levels: SoS Level (L1), Requirement Level (L2), and System Level (L3). Component elements at each level are a network representation of the level below. The SoS Level (L1) is comprised of the numerous, possibly interdependent, requirements (L2) needed to achieve a desired capability. Similarly, satisfaction of each requirement in the Requirement Level (L2) requires a number of possibly interdependent systems (L3). **Error! Reference source not found.** presents the description of the process modeled by the CEM.

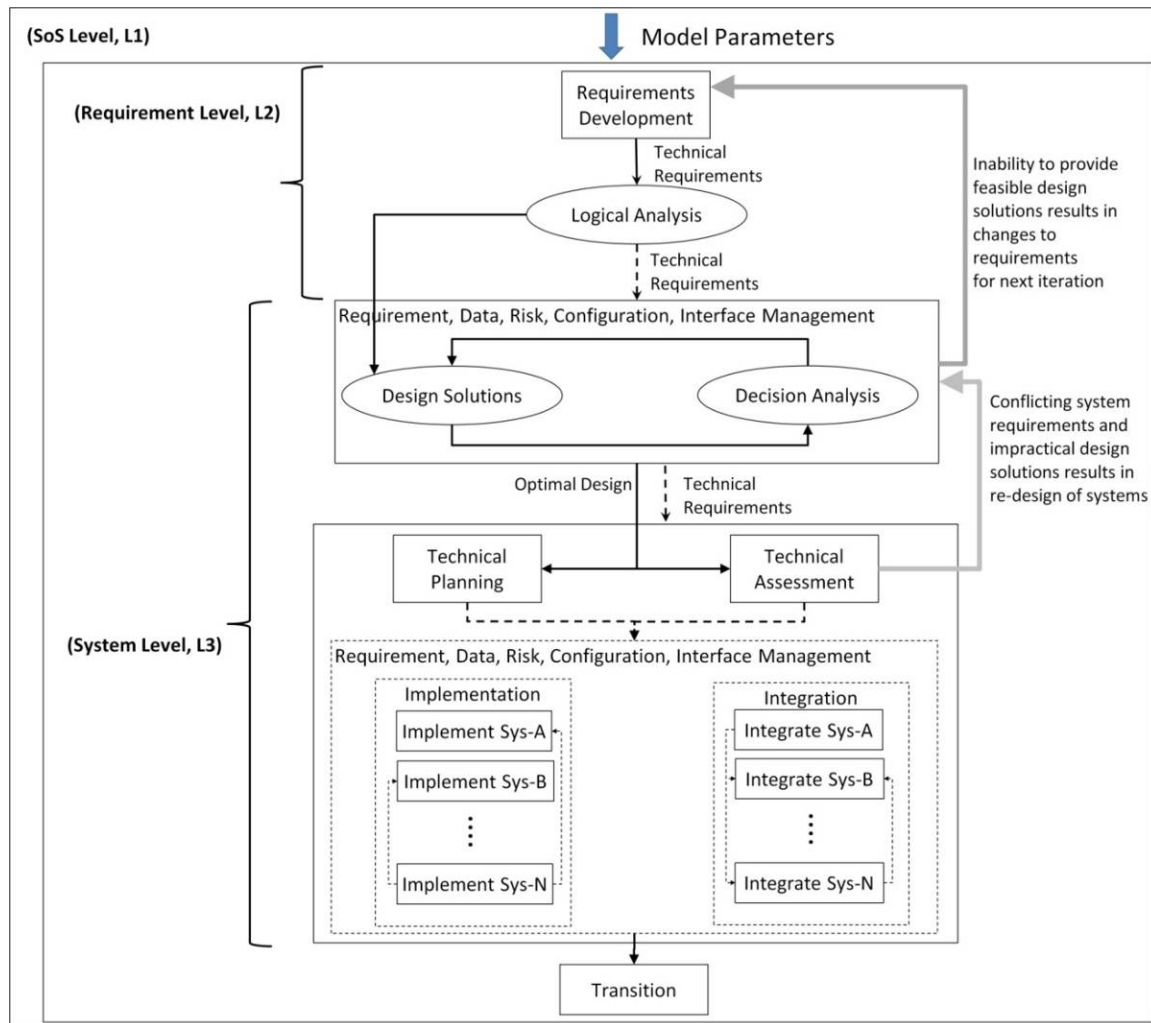


Figure 1. Conceptual Model of Acquisition Strategy based on SoSE Process
(described in DoD, 2008b)

At the Requirement Level (L2), *Requirements Development* contains the technical requirements of the SoS (provided externally). The technical requirements are then examined in *Logical Analysis* to check for interdependencies amongst the requirements. A check for inconsistencies amongst requirements is also performed. *Design Solution* development and *Decision Analysis* are the next processes, which belong to the System Level (L3). They produce the optimal design solution from the set of feasible solutions to meet the given requirements. The optimal design solution is based not only on the current set of requirements and solution alternatives but also takes into account all previous information available through requirements, risk, configuration, interface and data management processes. Because most acquisitions are multi-year projects involving many different parties, the overlap between the management processes, *Design Solution* and *Decision Analysis* processes, allows for greater tractability of decisions. It is at this stage that system interdependencies are identified. The optimal design solution obtained from this phase is then sent to the next stage: *Technology Planning* and *Technology Assessment*. In the event that an optimal or sub-optimal design solution to successfully implement the given requirements does not exist, the feedback loop to *Requirement Development* translates into a change in the technical requirements for the SoS. *Technology Planning* and *Technology*

Assessment are System Level (L3) scheduling processes that oversee the implementation, integration, verification and validation for all the component systems in the SoS.

Systems in the SoS are often dependent on other systems for either implementation, integration, or both. Disruptions during these stages of development in one of the systems result in time-lags in the acquisition process and to delays that propagate through the network of component systems impacting seemingly independent systems. For example, if the implementation of a system A is dependent on the Implementation of a system B—as could be the case for the development of an aircraft that depends on the specifications of a radar system—funding cuts to system B can result in development delays in system B but can also impact the development of system A. If, on the other hand, a third system C depends on system A, this could also be affected by the problems caused in system C due to funding cuts.

The *Implementation* and *Integration* Phases of component systems constitute the lowest level of detail modeled in the CEM. The design decisions made at earlier stages must be implemented and integrated in these phases to generate the final product of a program. **Error! Reference source not found.** presents an abstraction of the layered networks that result from the modeling of the acquisition process: systems are grouped to satisfy a requirement, and requirements are grouped to generate a capability.

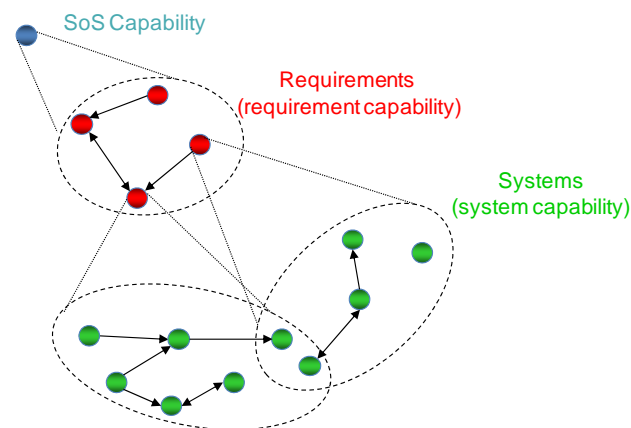


Figure 2. Layered Network Abstraction of Computational Exploratory Model

Systems can be independent, can satisfy several requirements, and can depend on other systems. The CEM simulates these layered relationships to capture the impacts that any changes—related to decision-making, policy, or development—in any of the component systems, requirements, and relationships between them have on the completion of a project. The exercise of the CEM described in this paper specifically targets complexities stemming from system risk, the interdependencies among systems, and the span-of-control of the SoS authority (if present). The next section will present the model dynamics that make possible the study of these complexities and will explore the design space of the SoS authority and tradeoffs between development risk and the number of systems and system interdependencies in an SoS.

Detailed Model Dynamics

The CEM operates as a discrete event simulator of the development process. Several challenges arise in developing a model for purposes of simulation and learning. Disruptions occur at various stages of development and are governed by the risk associated

with the project or individual systems. The CEM models risk associated with the implementation and integration of each component system as well as the risk due to the system interdependencies. Furthermore, systems and SoS engineers are often faced with the decision of using legacy assets to satisfy a given requirement or opt for the development of brand new ones. The CEM includes parameters such as *readiness-level* to differentiate between legacy assets/platforms, new systems, and partially implemented/integrated systems (i.e., systems under development) and to investigate the impact that the inclusion of such systems in the development of an SoS has on the success of a project. The next sub-sections describe the model details: parameters and inputs, Implementation and Integration dynamics, and the risk model.

Model Input Parameters

Error! Reference source not found. presents the input parameters and the remainder of this section expands and explains their role in the ECM.

Table 1. Input Parameters of Computational Exploratory Model

Parameter	Notation	Description
Requirement Level (L2)		
Requirement dependencies	D_{req}	Adjacency matrix that indicates requirement interdependencies
Risk profile	R_{req}	Probability of disruptions in <i>Requirement Development Phase</i>
Impact of disruptions	I_{req}	Time penalty when disruptions hit <i>Requirement Development Phase</i>
System Level (L3)		
System dependencies	D_{sys}	Adjacency matrix that indicates system interdependencies
Development pace of design	t_{des}	Increase in completion of <i>Design Solutions Phase</i>
Design risk profile	R_{des}	Probability of disruptions in <i>Design Solutions Phase</i>
Impact of design disruptions	I_{des}	Time penalty when disruptions hit <i>Design Solutions Phase</i>
Span-of-control	soc	Indicator of how <i>Implementation</i> and <i>Integration</i> are performed (sequentially or simultaneously)
System initial readiness-level	$m^0(i,r)$	Initial readiness-level of system i to satisfy requirement r (for <i>Implementation Phase</i>)
System risk profile	$R_{sys}(i,r)$	Probability of disruptions (during implementation) of system i when satisfying requirement r
Impact of disruptions	$I_{sys}(i)$	Time penalty when disruptions hit system i during Implementation/Integration
Implementation pace	$p_{imp}(i)$	Increase in readiness-level at each time step during implementation of system i
Integration pace	$p_{int}(i)$	Increase in completeness-level at each time step during integration of system i
Implementation start	$I_{imp}(i,j)$	Readiness-level of system j when <i>Implementation Phase</i> of dependent system i begins
Strength of dependency	$S(i,j)$	Strength of dependency of system i on system j

The requirement dependency matrix (D_{req}) indicates how the development and satisfaction of requirements depend on each other, which impacts the sequence in which requirements are developed and satisfied. For example, if Requirement A depends on Requirement B, then development of Requirement A begins when Requirement B has been satisfied. As requirements are developed, the risk profile (R_{req}) of *Requirement*

Development indicates the probability of disruptions at this stage in the development process. Disruptors signify a change in requirements or addition of new requirements. When a requirement is changed after the acquisition process has begun, it affects all subsequent processes and it causes a time delay (I_{req}) that is added to the project time. Every requirement that is implemented is fed into its own *Design Solution and Decision Analysis* (**Error! Reference source not found.**) process. The *Design Solution and Decision Analysis* processes feed into each other and the risk profile (R_{des}) indicates the probability of disruptions at each time-step during the completion of the stage with a value between 0 and 1. Any disruptions at this stage indicate that the design solution provided is not feasible and a time penalty (I_{des}) that indicates a re-design of the solution is incurred. If the solution fails in multiple consecutive time-steps, then the requirement is sent back to *Requirement Development* stage, otherwise the set of component systems and their user-defined parameters are sent to the *Technical Planning* and *Technical Assessment* (**Error! Reference source not found.**) processes based on the development-pace parameter of this stage.

Implementation Phase Dynamics

Technical Planning is the stage in which *Implementation* and *Integration* of component systems is performed. The *Implementation Phase* simulates the development of each system. The nature of candidate systems may range from legacy systems to off-the-shelf, plug-and-play products to custom-built, new systems. Development of a “brand new” SoS has been and will remain a rare occurrence. In their study on SoS, the United States Air Force (USAF) Scientific Advisory Board (Saunders et al., 2005) stated that one of the challenges in building an SoS is accounting for contributions and constraints of legacy assets. Similarly, the regular utilization of off-the-shelf component systems in both defense and civil programs contribute to cost and time savings but also introduce a different type of risk to the system development process (Constantine& Solak, 2010). These legacy systems may be used “as-is” or may need re-engineering to fulfill needs of the new program.

Here, we define *legacy systems* as systems that have been developed in the past to achieve a particular requirement, and *new systems* as not-yet-developed systems envisioned to satisfy a new requirement. When considering the use of legacy systems to meet a new requirement, the capability of these systems to satisfy the new requirement is not necessarily the same as their capability to meet the original requirement for which they were designed. Additionally, the risk associated with the modification of a legacy system and the risk associated with the development of a brand new system can be quite different. Legacy systems may, however, provide cost and/or time benefits if modifications are less severe than a new development, as is the case with new systems. To delineate systems in a meaningful way, we describe the spectrum of a system’s ability to satisfy a requirement in terms of its readiness-level.

System readiness-level, a concept proposed by Sauser et al. (2006), is a metric that incorporates the maturity levels of critical components and their readiness for integration (i.e. integration requirements of technologies). This is an extension of the widely used Technology Readiness Level (TRL), a metric that assesses the maturity level of a program’s technologies before system development begins (Department of Defense Directive 5000.2, 2005). While similar in spirit to the SRL metric proposed by Sauser, Verna, Ramirez-Marquez, and Gove (2006), readiness-level in the present work is defined in a different manner and with less detail. We define system readiness level as the readiness-level of a system i to satisfy requirement r , $m(i,r)$, with a value between 0 and 1. A system with a readiness-level of 1 is a fully developed system that can provide a certain level of capability.



The dynamic model starts the *Implementation Phase* of a system from its initial readiness-level and simulates its development/implementation until it reaches a readiness-level of 1. An initial readiness-level of 0 indicates a brand new system that must be developed from scratch, while a system with an initial readiness-level greater than 0 indicates a legacy system that is partially developed to satisfy a requirement r , but needs further development to reach a readiness-level of 1. In general, careful research of a candidate system i will determine its initial readiness-level to satisfy a requirement r , and, therefore, the amount of development necessary to achieve a readiness-level of 1.0.

The CEM simulates the *Implementation Phase* as a series of time steps in which a pre-determined increment of readiness ($p_{imp}(i)$) is gained at each time-step of each system i , or lost if a disruption occurs (according to the system risk profile of system i in satisfying requirement r , $R_{sys}(i,r)$). This is clearly a gross simplification of the actual development process for a system; however, it adequately serves the purposes of the research, which is focused on the interdependencies between systems to develop an SoS capability and aims to capture the impact of disruptions on the development process. Accurate modeling of the *Implementation Phase* would increase the accuracy of the model for a particular application, but it would not change the nature of the observed results.

Representation of Risk

The risk associated with the development of a system is a function of its inherent characteristics (technology, funding, and complexity levels) and on risk levels of the systems on which it depends. The former may be estimated via a variety of analysis techniques that examine a system in detail, but the latter requires knowledge of system interdependencies which can be numerous, complicated, and often opaque. Developmental interdependencies of SoS create layered networks that often span among a hierarchy of levels (DeLaurentis et al., 2008; Butler, 2001; Ayyalasomayajula, DeLaurentis, Moore & Glickman, 2008; Kotegawa, DeLaurentis, Sengstacken & Han, 2008). The complexity of these networks often hides many of the otherwise explicit consequences of risk. Depending on the network topology characteristics, disruptions to one of the critical nodes or links in the network can propagate through the network and result in degradation to seemingly distant nodes (Huang, Behara & Hu, 2008).

In this study, we express risk as a density function that describes the probability of a disruption occurring at any time during the system development. We concentrate on the *Implementation* and *Integration Phase* as the development stage in which disruptions occur. Here, inherent risk is the probability of disruptions due to the development characteristics of the subject system, e.g., technology readiness-level, funding, politics, etc. Risk due to interdependencies, on the other hand, is the probability of disruptions during the *Implementation Phase* of a system due to disruption in the system on which the system of interest depends. This is essentially the conditional probability of a disruption given that another system has a disruption.

This study assumes that the inherent risk of a system i in satisfying requirement r , $R_{sys}(i,r)$, is solely a function of its readiness-level, $m(i,r)$. While a somewhat simplified definition, expressing risk as a function of a system's readiness-level is logical. Recall that readiness-level is a metric that describes the necessary development of a system to satisfy a given requirement. Therefore, risk changes as the readiness-level of a system increases. Equation 1 introduces a relationship between a system's readiness-level and risk (probability of disruption).



$$R_{sys}(i, r) = \alpha_i (1 - m(i, r)^{\beta_i})$$

Equation 1

In this relationship, α_i (with a value between 0 and 1) is parameter that indicates the upper bound value of risk for system i (i.e., producing maximum probability of disruption), while β_i is a shape parameter that indicates how quickly risk changes as a function of readiness-level. This formulation implies that risk is highest at the early stages of development (e.g., low readiness-levels) and it decreases (at different rates depending on the value of the β_i parameter) as development progresses. For instance, when a system i has a readiness-level of 0.0—it is a brand new system—the probability of disruptions during development will be highest, and it will have a value α_i . However, when the system has a readiness-level of 1.0, the probability of disruptions will be 0. System inherent-risk is implemented in the CEM by using a uniform random distribution to select a value between 0 and 1 at each time-step of the *Implementation* or *Integration Phase* and passing it into a binary channel to see if the number is smaller or greater than the probability of disruption defined by $R_{sys}(i, j)$. This determines if a disruption occurs or not.

When all systems are independent, identification of the system with highest risk is trivial (e.g., system that, on average, will contribute more to delays in completion time). However, when systems are interdependent, systems that otherwise have a low inherent risk can be greatly impacted by disturbances because of the transmission of risk from other systems. Systems are impacted by nearest neighbors (those systems on which they directly depend; first-order dependencies) and by systems that impact those nearest neighbors (higher-order dependencies).

The CEM models risk due to interdependencies in terms of the dependency strength between two given systems. Dependency strength, $S(i, j)$, is an input parameter that takes values between 0 and 1 and is defined as the conditional probability (uniform random probability) that system i has a disruption given that system j (on which system i depends) has a disruption. Risk due to interdependencies is, therefore, a function of the readiness-level of the dependent-upon system as well as the strength of that dependency. A notional example of a simple SoS is utilized here to present these features of the CEM (**Error! Reference source not found.**).

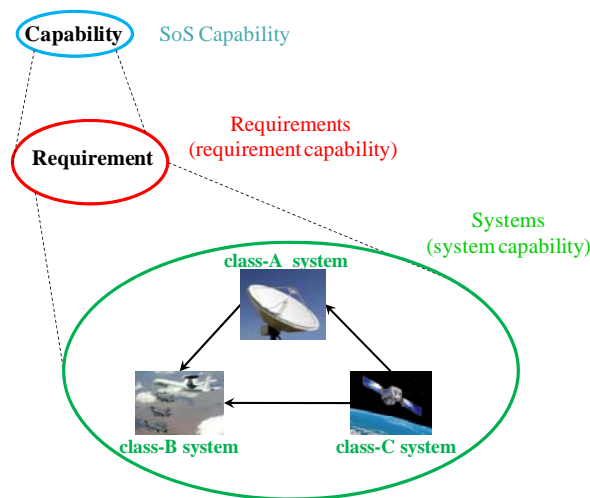


Figure 3. Layered Network Structure of Example SoS

Each system in this simple SoS network serves a role and provides a certain level of capability in order to satisfy some requirement. The links between systems indicate interdependencies among systems. The arrows indicate the directionality of dependence, including the case of mutual dependence. Mane and DeLaurentis (2009) contain more detailed information on the CEM structure. For this example, **Error! Reference source not found.** presents the implementation history of this three-system SoS with a risk profile that has α_i and β_i values of 0.2 and 4, respectively, and two different levels of interdependency strength, $S(i,j)$.

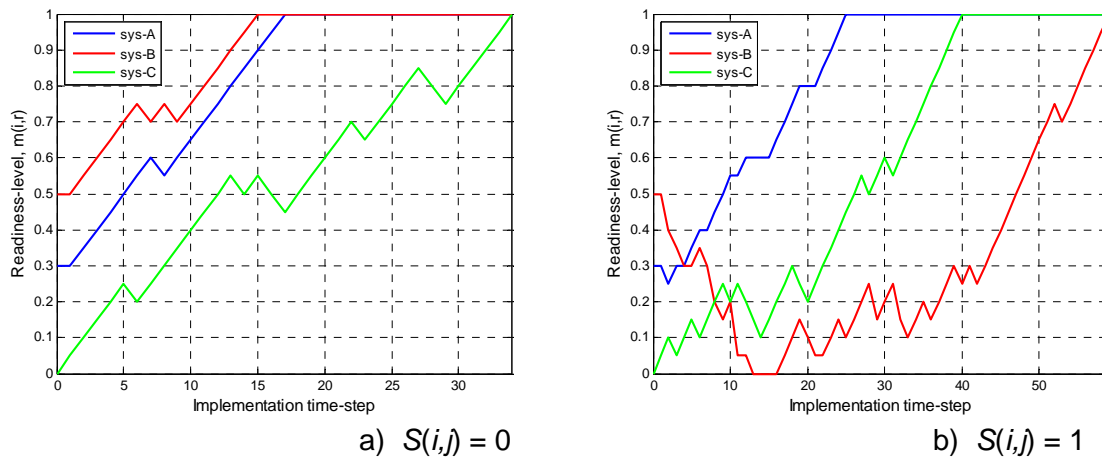


Figure 4. Implementation Phase History for Example Problem

Each system has a different initial readiness-level—system-A of 0.3, system-B of 0.5, and system-C of 0. Recall that an initial readiness-level greater than zero indicates a legacy system that must be further developed to achieve a readiness level of 1 to satisfy a given requirement. The model assumes that the readiness-level of a system can reduce to below initial readiness-level value. This is reasonable since inherent disruptions or disruptions due to interdependencies can result in modifications to subsystems that were not previously considered (i.e., unforeseen technology limitations of a system may require redesign of a dependent system). In **Error! Reference source not found.a**, all systems are independent (dependency strength of zero). The occasional set-backs in the readiness-level of each system are due to disruptions stemming from the inherent system risk. In **Error! Reference source not found.b**, on the other hand, dependency strength is highest (with a value of one). Recall that dependency strength indicates the probability of disruption on the dependent system given that the system on which it depends has a disruption. When the dependency strength is one, a disruption in a given system is always propagated to the dependent systems. For example, disruptions in the development of system-C propagate to system-A with probability 1 and disruptions in the development of system-A propagate to system-B with probability 1. Note, for instance, that there is a reduction in readiness-level in the development of the system-B every time that there is a reduction in readiness-level during the development of system-A or system-C (on which system-B depends). The candidate systems for a desired capability can, in general, have different levels of dependency strengths. **Error! Reference source not found.** presents a sensitivity of development time for this example problem on the value of dependency strength.

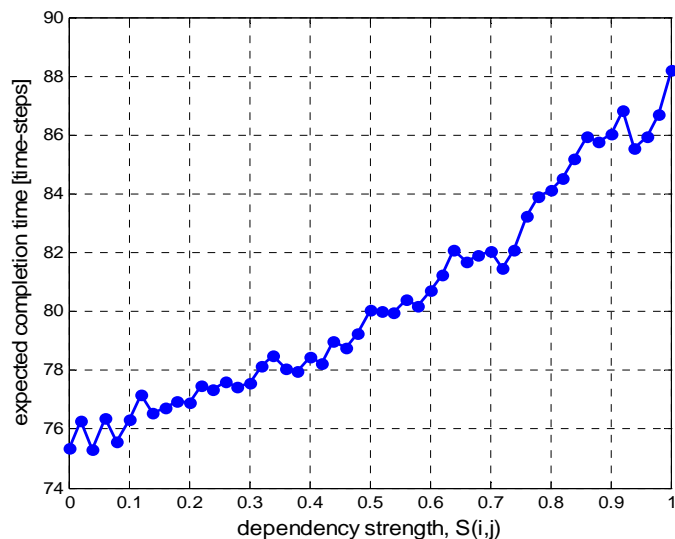


Figure 5. Impact of Dependency Strength on Completion Time for Example Problem

As expected, higher dependency strength means higher development time. In this example, the number of systems and interdependencies is invariable, and the increase in development time can be different for a different family of constituent systems. When considering the development of different families of systems that can provide a desired capability, the characteristics of interdependencies between component systems can have a large impact on the decision to pursue development of a certain alternative. Quantifying the impact that such characteristics have on the development process can aid decision-makers in selecting the most promising alternative.

Impact of Risk and System Interdependencies

Quantifying risk is a complicated function of the individual system characteristics as well as the interdependencies between systems. The combinations of systems that can achieve a given capability-level can be numerous. Depending on the selection of the constituent systems, the completion time of a project can vary greatly due to the number of constituent systems, their interdependencies, and risk profiles. As these families of systems get larger, it becomes more difficult to quantify the impact that each system and system-characteristic has on the success of a project. For instance, a three-system solution may appear to be preferable to a ten-system solution; but the interactions between the three systems can result in disruption propagation that greatly impacts the timely completion of the project. System interdependencies and their characteristics can impact the completion time of a project by affecting the way in which disruption propagate. In this section, we demonstrate the impact that system-inherent risk and the strength of interdependencies between component systems can have on the timely completion of a project. Furthermore, we show and quantify how different families of systems that can provide the same set of capabilities can have greatly differing development histories.

Interdependency Strength and Inherent Risk

For this investigation, we assume that in order to achieve some capability a family of three classes of systems has been identified; for instance, a class-A system can be a land-based radar or an airborne radar; a class-B system can be a large transport aircraft, a mid-

size aircraft, or a small aircraft. Each of these classes of systems provides a certain capability that is required to achieve a global capability of the SoS. The design authority must decide which constituent system to select for each system-class. A notional example of a simple SoS is utilized here (**Error! Reference source not found.**).

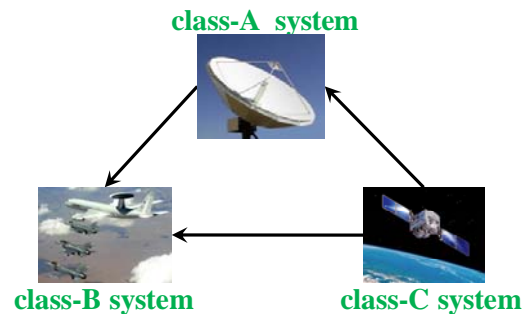


Figure 6. Interdependencies of Notional SoS

The links between systems indicate interdependencies among systems. For instance, development of a class-B system must rely on information about the development and capabilities of a class-A system in order to continue development. Similarly, development of a class-A system needs information from a class-C system. Different systems are available to designers or systems engineers for each system-class. Each candidate system can have different risk characteristics as well as different interdependency characteristics. If we assume that the systems engineer has identified these characteristics for each candidate system, then we can use the CEM to simulate the development process when different combinations of these candidate systems are considered and identify the family of systems that results in the lowest expected completion time. The strength of the CEM is in its ability to aggregate the individual system characteristics and quantify the SoS-level performance (with respect to development time) of a family of candidate systems.

Error! Reference source not found. presents results in which the expected implementation time of a family of candidate systems is measured against the inherent risk of individual systems and their interdependency strengths.

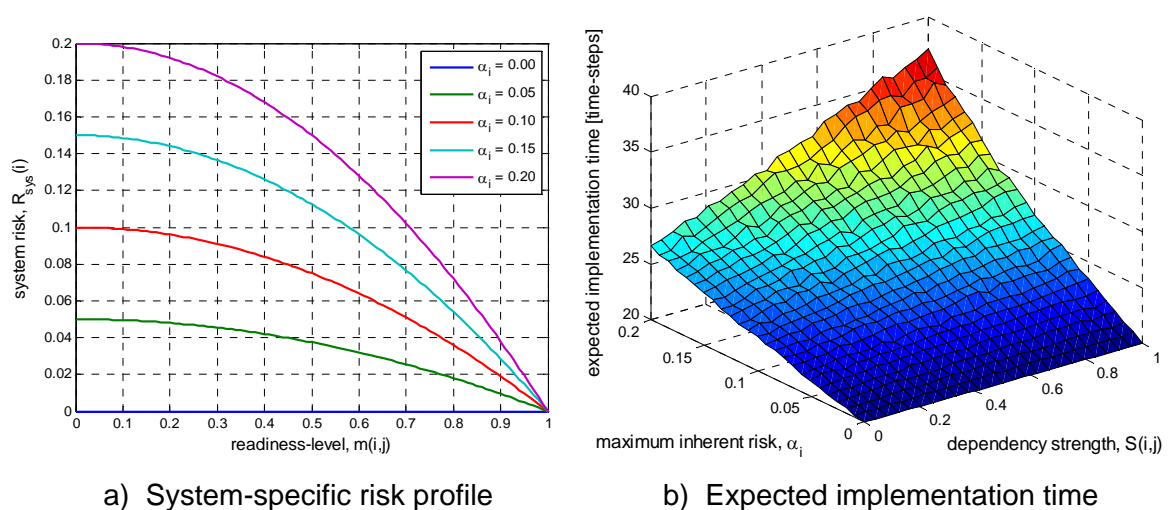


Figure 7. Impact of Risk Due to Interdependencies on Implementation Time

We assume here that all candidate systems will have the same risk profile and all interdependencies will have the same strength. **Error! Reference source not found.a** shows the inherent system risk, $R_{sys}(i,r)$, as a function of system readiness-level, $m(i,r)$, for five different risk profiles (five different α_i values and a fixed β_i parameter of 2). The value of α_i indicates the maximum inherent risk of a system, according to Equation 1. The assumption here is that risk is highest in the earlier stages of development and that it decreases as development progresses. The results in **Error! Reference source not found.b** present the expected implementation time when families of systems with different combinations of inherent risk profile and dependency strengths are considered. Each point on the surface indicates a family of candidate systems with a given combination of maximum inherent risk and dependency strengths. For instance, a solution that entails systems with a maximum inherent risk of zero and dependency strength of zero (e.g., independent systems with no development risk) will have an expected implementation time of 20 time units. The three systems are developed simultaneously but have no impact on each other's development. The trends in **Error! Reference source not found.b** show that the impact on implementation time of families of systems that have strong interdependencies is larger than when the systems have high inherent risk but low dependency strengths (e.g., the increase in implementation time is smaller as inherent risk increases than when the strength of dependencies increases).

This investigation quantifies the impact that system interdependencies have on the implementation time of a project. The results presented here point out the importance of interdependencies in the development process. This type of analysis can prove useful to an SoS authority when selecting potential component systems as a part of a family of systems or SoS to satisfy a given requirement and achieve a desired capability.

This simple example considers families of systems comprised of three constituent systems. Different candidate families of systems, however, can have differing number of constituent systems that can provide different system-capabilities to achieve the desired SoS capability. Similarly, risk profile and interdependency characteristics of the constituent systems can result in different disruption propagation and different development solutions.

Comparison of Alternatives

Given a set of alternative means to satisfy a requirement, an SoS authority (in conjunction with systems engineers) must determine the best network of systems to develop and acquire. The number of systems alone may not be a good indicator of the complexity of a system and the eventual developmental success. The risk profile of systems as well as the number and strength of system interdependencies play an important role that often hamper understanding of the impact of decisions. For instance, an SoS that is comprised of three constituent systems may appear more likely to succeed than an SoS comprised of five systems. However, the number and strength of interdependencies between the five systems may be such that the expected completion time of this SoS is lower than the expected completion time of the three-system SoS. The three-system example in the previous section showed that the strength of dependencies plays an important role in the timely completion of an SoS project. Here we use the CEM to investigate the impact that network characteristics (number of systems, number of dependencies, and strength of dependencies) have on the completion time of an SoS project. We compare the developmental time of two example SoSs comprised of three and five constituent systems (**Error! Reference source not found.**).

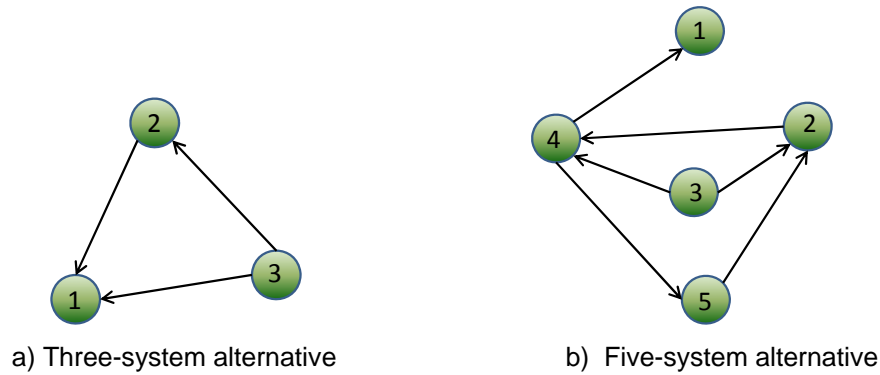


Figure 8. Alternative Families of Systems

The three-system network is the same network with three interdependencies as the one presented in **Error! Reference source not found.**. The new, five-system network is clearly a larger SoS with more systems and six interdependencies. As in the previous section, different candidate systems are available to provide the required capability level. The systems engineer would like to quantify the expected implementation time of each combination of systems for the three-system and the five-systems options. Via a Monte Carlo simulation of 500 samples, we are able to compute the expected implementation time of the five-system network—we previously did the same for the three-system network. **Error! Reference source not found.** presents this result for the different combinations of inherent system risk and dependency strengths.

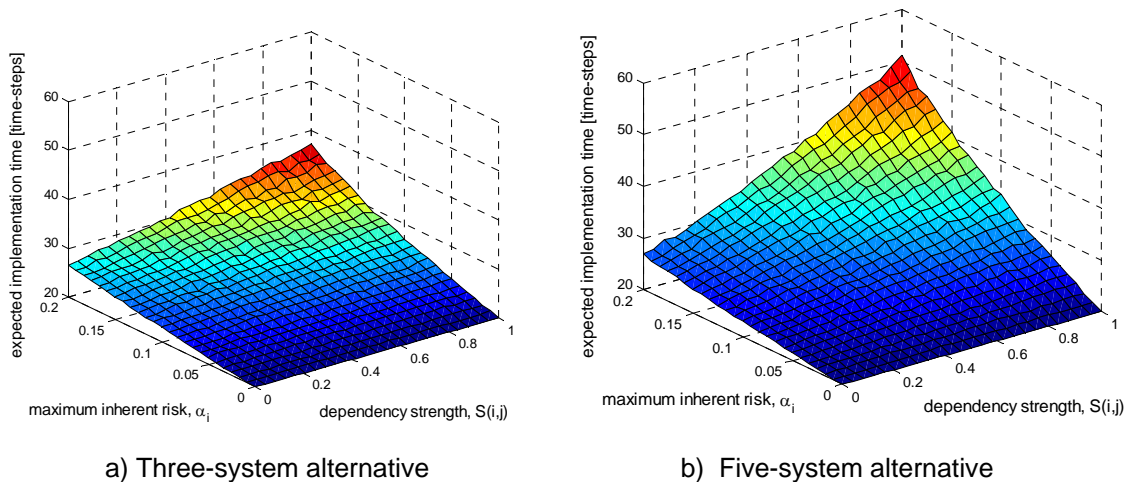


Figure 9. Expected Implementation Time of Alternatives

Error! Reference source not found.a presents the expected implementation time of the three-system option (the same as **Error! Reference source not found.**b), while **Error! Reference source not found.**b presents the expected implementation time of the five-system network. As in the previous analysis, these results indicate the expected implementation time of candidate component systems that have differing levels of inherent risk and interdependency strengths. The trends in the expected implementation time of the five-system option are larger than those of the three-system option. This is expected because the former has more systems as well as more interdependencies. Recall, however, that each point in these charts represents a candidate family of systems and one can see

that the expected implementation times of some five-system alternatives are lower than some three-system alternatives. To show this more clearly, **Error! Reference source not found.** presents the expected completion times when the inherent system risk of all candidate systems is highest ($\alpha = 0.2$).

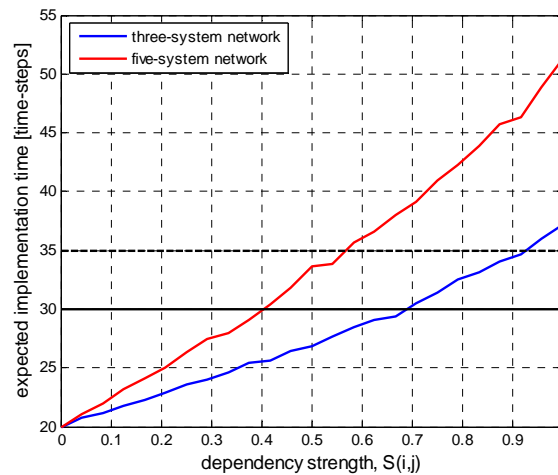


Figure 10. Expected Implementation Time of Sample Results

As previously mentioned, the implementation time of the five-system alternatives is always higher than the three-system alternatives. However, if the dependency strength between the systems in the three-system alternative has a value of 1, then the expected implementation time of this alternative will be 37 time units; if the dependency strength between the systems in the five-system is not as strong, say with a value of 0.4, then the expected implementation time will be 30 time units. Therefore, depending on the strength of the interdependencies between the constituent systems, a family of systems can be a better (lower expected implementation time) alternative. By simulating the development process of different alternatives via the CEM, it is possible to quantify the impact of system specific risk, the risk due to interdependencies, and the propagation of disruptions to compare different alternative solutions that can provide a desired level of capability.

Analytical Approach to Measure Delay Propagation

Additional complexity in the model, carefully selected, will likely increase the efficacy of the CEM. However, as a simulation-based approach, it too has limitations. Therefore, in conjunction with the further development of the CEM, the authors are also developing an analytical approach that captures the characteristics of a network that results from the developmental interdependencies of systems. This is an approach that uses a network-level metric to treat the same complexities via computations on conditional probabilities that relate the transmission of risk in networks of interdependent systems. This provides means to compare networks in their ability to arrest the propagation of delays caused by random disturbances and can be used as a figure of merit when designing SoS architectures that aim to achieve some desired capability. While typical networks like the World Wide Web, social networks, and communication networks are a result of evolution, some networks of military systems created for particular purposes can be designed. The ability to quantify the performance of SoS networks enables comparison of networks, and ultimately the design of superior SoS networks that optimize that performance.

The proposed approach to measure the performance of networks in their ability to arrest the propagation of delays is based on the “lost miner problem” (Ross, 2007). In this example problem, a miner is lost in a cave inside a mine and there are four tunnels that lead out of the cave, but only one leads out of the mine (**Error! Reference source not found.**).

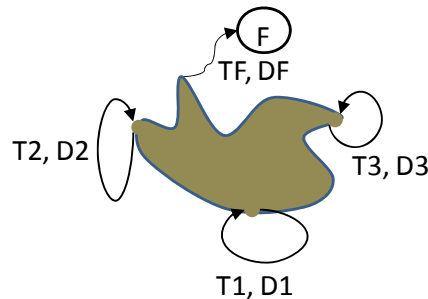


Figure 11. Lost-miner Problem

The miner can choose to enter a tunnel T_i with probability $P(T_i)$ and has no memory of his previous choice. If the miner chooses tunnel T_1 , then he wanders in the tunnel for D_1 days and returns to the cave, where he must decide which tunnel to enter next. If he chooses tunnel T_2 or T_3 , then he wanders in the tunnel for D_2 or D_3 days, respectively, only to return to the cave. If he chooses tunnel T_F , then he is free, instantly. The question the problem poses is: What is the expected time until the miner reaches freedom (e.g., the expected duration of the miner’s stay in the mine)?

Following this reasoning, we can describe the delay propagation in the system development process in a similar manner. We describe a network of systems in terms of the number of systems (caves), the number and direction of their dependencies (tunnels), and the characteristics of the interdependencies (probability of choosing a given tunnel), e.g., probability of passing on a disruption and the impact of the disruption. The simple three-system network below is used to describe the proposed approach.

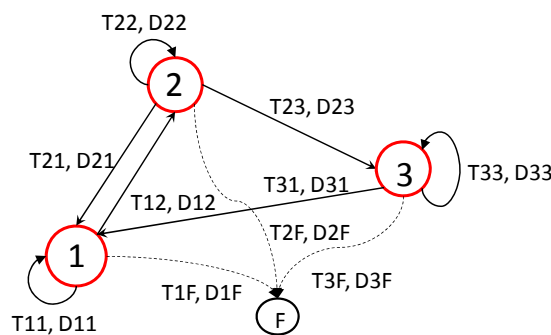


Figure 12. Example Systems Development Network

Each node represents a system that is under development (i.e., aircraft, missile, radio) to achieve some capability. The links indicate interdependencies between the systems as well as the strength of those interdependencies. For instance, system-1 depends on system-3 because information from system-3 is needed to continue development of system-1. T_{ij} represents the conditional probability that a disruption in the development of system i will impact development of system j and D_{ij} represents the impact of a disruption (delay) on system i that propagates to system j . These two quantities represent the strength of the dependency between system i and system j . Two systems can

be strongly dependent if the probability of a disruption propagating from one system to the other is high, or if the delay experienced by one system because of a disruption in the development of the other is large. Node F is a sink that represent the arrest of an event and its propagation in the network. In this setting, a disruption can be seen as an event that travels from system to system causing developmental delays until it exits the system/network (via node F). This is similar to the “lost miner problem,” in which the miner chooses tunnels until he reaches freedom.

In system development, disruptions can be a result of funding decisions, political environment, technological setbacks, etc. For example, system-1 can be faced with budget cuts and the program manager must reduce funding to one of the subsystems that comprise system-1. Depending on the magnitude of the reduction in funding, this can have no impact in the development of system-1 with probability T_{1F} , and nothing is affected; it can cause a delay of D_{11} days with probability T_{11} in the development of system-1 that is not large enough to impact interdependent systems; or it can result in a delay of D_{12} days with probability T_{12} that impacts development of system-2. Additionally, the delay in the development of system-2 can cause further problems that delay its development by D_{22} days with probability T_{22} ; it can cause a delay of D_{23} days with probability T_{23} that creates a problem in the development of system-3; or a delay of D_{21} days with probability T_{21} that impacts system-1; or, conversely, the problem is not large enough to cause any delays with probability T_{2F} , and the propagation of delay in the network is arrested.

Depending on the strength of the dependencies between systems and the magnitude of disruptions, delays can propagate and accumulate in a network. Hence, networks with different number of systems, interdependencies, and strength of interdependencies will perform differently when faced with random disruptions. The analytical approach now under pursuit may be able to estimate the expected accumulation of delays as a function of these network characteristics. The network-level metric can enable the design of networks that minimize expected delay whenever random events hit the development process of individual systems.

Conclusions

The development of complex systems (and SoS) is beset by risk. Risk analyses of individual systems can explain the threats and opportunities of systems, but do not capture the impact that disruptions to individual systems have at the enterprise level, where multiple systems—explicitly or implicitly interdependent—collaborate to achieve various capabilities. An understanding of risk in the development and acquisition process and its cascading effects is crucial to identifying means to exploit opportunities, as well as mitigate, transfer, or avoid disruptions.

These research efforts center on the ongoing development of a Computational Exploratory Model that is based on the processes in the *SoS-SE Guidebook* and that estimates time to complete an SoS integration. The extensions to the model in this paper capture the impact of individual system risk and number and strength of system interdependencies on the propagation of developmental disruptions and, ultimately, the timely completion of a project. In particular, the present work examined changes in the systems interdependencies and system risk profiles (i.e., different inherent risk profiles for different systems and different dependency strength) when alternative systems are considered for satisfying a given requirement and providing a certain capability. Examples of alternative families of systems comprised of a different number of constituent systems showed that the number of constituent systems and their risk profiles are insufficient to

quantify the development performance of SoS. The sample analyses presented here showed that these characteristics, coupled with the interdependency characteristics of a family of systems, can result in expected implementation times that are not easily foreseen.

When coupled with the theoretical basis of delay propagation and a network-level metric that describes the expected delay in a family of systems the methodology presented here can improve/facilitate the decision-making process of systems engineers and system integration as well as provide a means to design system architectures that aim to minimize delay propagation and development time.

Acknowledgments

The authors acknowledge the support for this work from a grant from the Naval Postgraduate School's Acquisition Research Program (N00244-09-1-0011).

References

- Ayyalasomayajula, S., DeLaurentis, D.A., Moore, G.E., & Glickman, L.T. (2008, October). A network model of H5N1 Avian Influenza transmission dynamics in domestic cats. *Zoonoses and Public Health*, 55(8-10), 497-506.
- Brown, M., & Flowe, R. (2005). Joint capabilities and systems of systems solutions: A multidimensional approach to understanding cost drivers. *Defense Acquisition Review Journal*, 12(2), 138-154.
- Butler, J.T. (2001, April). *UAVs and ISR sensor technology (Research Report)*. Air Command and Staff College, Air University, AU/ACSC/033/2001-04. Retrieved January 26, 2010, from <https://research.maxwell.af.mil/papers/ay2001/acsc/01-033.pdf>
- Charles, P., & Turner, P. (2004). Capabilities based acquisition...From theory to reality. *CHIPS Magazine*. Retrieved January 21, 2010, from http://www.chips.navy.mil/archives/04_summer/web_pages/GEMINI.htm
- Committee on Systems Integration for Project Constellation. (2004). *Systems integration for project constellation*. Retrieved January 15, 2009, from The National Academies, <http://nap.edu/html/proj-constellation/ltr-rep.pdf>
- Constantine, J., & Solak, S. (2010). SysML modeling of off-the-shelf-option acquisition for risk mitigation in military programs. *Systems Engineering*, 13(1), 80-94.
- DeLaurentis, D., Han, E-P., & Kotegawa, T. (2008). Network—Theoretic approach for analyzing connectivity in air transportation networks. *AIAA Journal of Aircraft*, 45(5), 1669-1679.
- DoD. (2008a). *Defense acquisition guidebook*. Retrieved April 4, 2008, from <https://akss.dau.mil/dag/>
- DoD. (2008b). *Systems engineering guide for system-of-systems*. Retrieved March 2, 2008, from http://www.acq.osd.mil/sse/ssa/initiat_sos-se.html
- Durkac, L.M. (2005, March). *Joint mission capability packages: The future of joint combat*. Presented at the 10th International Command and Control Research and Technology Symposium. Retrieved January 25, 2010, from online database, http://www.dodccrp.org/events/10th_ICCRTS/CD/papers/063.pdf
- Francis, P. (2007, April). *Defense acquisitions: Missile Defense Agency's flexibility reduces transparency of program cost*. Testimony before the Subcommittee on Defense, Committee on Appropriations, US House of Representatives. Washington, DC: Government Accountability Office. Retrieved January 27, 2010, from www.gao.gov/new.items/d07799t.pdf



- Gilmore, M.J. (2006, April). *The Army's Future Combat System Program*. CBO Testimony. Retrieved March 3, 2010, from <http://www.cbo.gov/ftpdocs/71xx/doc7122/04-04-FutureCombatSystems.pdf>
- Huang, C.D., Behara, R.S., & Hu, Q. (2008). Managing risk propagation in extended enterprise networks. *IT Professional*, 10(4), 14-19.
- Kotegawa, T., DeLaurentis, D.A., Sengstacken, A., & Han, E-p. (2008, September). *Utilization of network theory for the enhancement of ATO air route forecast*. Anchorage, AK: 8th AIAA Aviation Technology, Integration, and Operations.
- Maier, M. (1998). Architecting principles for system-of-systems. *Systems Engineering*, 1(4), 267-284.
- Mane, M., & DeLaurentis, D. (2009, May). Acquisition management for systems of systems: Exploratory model development and experimentation. In *Proceedings of the Sixth Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School.
- NextGen Integration and Implementation Office. (2009). *NextGen implementation plan 2009*. Washington, DC: Federal Aviation Administration. Retrieved September 25, 2009, from http://www.faa.gov/about/initiatives/nextgen/media/NGIP_0130.pdf.
- Ross, M.S. (2007). *Introduction to probability models* (9th ed.). Burlington, MA: Academic Press.
- Rouse, W. (2001). Complex engineered, organizational and natural systems. *Systems Engineering*, 10(3), 260-271.
- Saunders, T., et al. (2005). *Report on system-of-systems engineering for Air Force capability development* (SAB-TR-05-04). Washington, DC: USAF Scientific Advisory Board.
- Sauser, B., Verna, D., Ramirez-Marquez, J., & Gove, R. (2006). From TRL to SRL: The concept of systems readiness levels. In *Proceedings of the Conference on Systems Engineering Research*. Los Angeles, CA: CSER.
- Spacy, W.L., II. (2004). Capability-based acquisition in the Defense Agency and implications for Department of Defense acquisition. *Journal of Contract Management*, 10-19.
- USD(AT&L). (2005). *The defense acquisition system* (DoDD 5000.2). Chapter 4: Systems Engineering. Washington, DC: Author



Panel #20 – Managing Services Procurement in DoD

Thursday, May 13, 2010

1:45 p.m. –
3:15 p.m.

Chair: Jeffrey P. Parsons, Executive Director, Army Contracting Command,
US Army Material Command

The Changing Face of Procurement Policy—An Innovative Approach to Competing Requirements

Brian Johnson, Department of Veteran Affairs

Services Supply Chain in the Department of Defense: Comparison and Analysis of Army, Navy, and Air Force Empirical Survey Results

Aruna Apte, Uday Apte and Rene Rendon, Naval Postgraduate School

General Services Administration Streamlines the Procurement of Construction Services

Jeffory Meyer and Stephanie Witt, General Services Administration;
Dean Kashiwagi and Jacob Kashiwagi, Arizona State University



The Changing Face of Procurement Policy—An Innovative Approach to Competing Requirements

Brian Johnson—Brian Johnson is an intern with the inaugural Contract Specialist Intern class at the VAAA. Mr. Johnson has over 10 years of management experience in industry and military. He acquired extensive training and experience, including various assignments throughout the VA. Prior to joining the VA, he held the position of Director, Medical Assisting/Medical Administrative Specialist Programs at ACT College. In addition, he continued to excel as an instructor at both the associate degree and career diploma levels. Mr. Johnson's military experience includes duties in combat arms, military intelligence, and as a combat medic at both domestic and overseas assignments.

Abstract

Everyday, contracting officers must procure the goods and services for the Federal Government, while balancing competing and oftentimes conflicting demands. Over time a number of laws, regulations and guidance have been put in place to direct the actions of the Government as it interacts with the private sector in this country. As a result of our political process and the method in which laws are created, some of these requirements have conflicted with other guidance already in place within the Federal acquisition process. Several examples are present throughout our history, and in current regulations, to illustrate these forces that the contracting officer must manage. One example is the balancing act required between full and open competition, under *Competition in Contracting Act (CICA)* and the direction given by the *Federal Acquisition Regulation (FAR)* to set aside requirements for different types of small disadvantaged businesses. However, through innovative strategies, the creative contracting officer can balance these requirements while providing the best value for the Government. This paper frames the context of how these competing demands have evolved and how the government can balance these different needs and responsibilities.

Introduction

Every day across the government a group of individuals known as Contracting Officers navigate the minefield of procurement policy to acquire the goods and services required for the Government to function. On the other side of this field are the vendors, also known as contractors, competing among each other to “win” the business of the government. As with any competition, certain rules apply to “the game.” The origins of this game, for the government, go back to the early days of our country. As the environment the Government operates in changed, the methods of procuring goods and services have evolved. One cannot fully understand the Government procurement process without a little background in the political development of this country. As times change, new laws are passed affecting the procurement process. A number of these laws do have some political influence and address the perceived issues of the time.

The contracting game is not unique to the Federal Government. State governments and private industries buy goods and services from other companies in order to operate more efficiently. As the business world moves faster and becomes more global, companies have had to focus on their core strengths and outsource certain functions and components to remain competitive. As such, the Government did not invent contracting. However, just because the private industry does contracting, does not mean it is identical to Government contracting. Often when a company first enters Government contracting, it stumbles through



the process and make certain mistakes, even if the company has conducted private contracts for some time. Even though the Government tries to model a number of best practices from industry, different rules do apply.

When I first lived in England, after growing up in the United States, I found out two things right away. First, while we both speak English, it is not the same language. Second, if you describe Cricket to someone, it may sound like baseball, but it is a completely different game. In Cricket, a pitcher throws to a batter that tries to hit the ball. If he does hit it, he runs between bases, until someone from the other team gets the ball and tries to throw him out. After hearing this description, you may think you are ready to watch a game and know what to expect. However, with the subtle differences in the rules, most people watching this sport for the first time are as lost as I was trying to follow the game. One must have a strong understanding of the rules in order to follow or play the game. Government contracting is similar; you must know all the rules of the game in order to play successfully.

As long as everyone plays by the rules, the result is a fair competition. Unfortunately, it would appear, the rules in the Government contracting game often change without thoroughly looking at the impact on other rules. In order to understand how this happens, we must first examine the history of our Government.

A New Type of Government

To refresh from high school Government class, the founders of the United States Government mostly came from countries that had governments that exercised complete power over their citizens. Whether it was a monarchy, dictatorship, or some other form of totalitarian rule, the people had little say in how the government exercised its power over the people. As such, the founders of this country were greatly concerned about the Government getting too powerfully and imposing its will on the people. They did not want this new Government to become a version of the governments from which they just broke away.

The Constitution of the United States is the governing document for our country, often referred to as “the law of the land.” The founders designed it to spell out the structure of the federal government, and its powers in relation to the state governments. In developing the Constitution, the Founders felt very strongly that it needed to address the government’s interactions with the people and set limits on the power it could exercise over the people. The basic premise is that the government derives all its power from the people and must represent the people that it serves.

The first sections of the Constitution of the United States address the structure of the Government and the powers of each branch, to include methods to balance those powers. This section specifically states that the government derives its power from the people it governs. In addition, any powers that are not explicitly given to the Federal Government by the Constitution remain the power of the individual states. The Founders saw a need for a Federal Government to unite the states and act on certain matters that affect all the states. However, they were cautious of a central government abusing its power over the people. The first ten amendments, also known as the Bill of Rights, focuses completely on how the Government interacts with the people of the United States and restricts the powers the Government can exercise over its people.

To further restrict this power and ensure the government did not turn into an “elected monarchy,” the Federal Government was divided into three branches. Each branch has its role and area in which it has power to act. These powers put “checks and balances” on the other areas of Government for the purpose of ensuring one branch cannot get too powerful



and impose its power on the people, unchallenged. In the simplest definition, each branch has a unique duty and power. The Legislative Branch is responsible for passing all Federal laws. Once passed, the President, Executive Branch, “signs the legislation into law.” The Executive Branch is responsible for enforcing the laws and managing how the government operates. The White House and the Office of the President cannot accomplish all the responsibilities of running the different aspects of government by itself. To accomplish these responsibilities, the Executive branch is further broken down into specialty areas through the establishment of Executive Agencies and Administrations. The Executive Branch manages through directives and Executive Orders to its Administrations. The Judicial Branch is responsible for interpreting the laws and stating how the laws apply to a specific situation. The Judiciary also passes judgment when someone files a complaint that a law or regulation was violated. This judgment can include indemnification of an injured party. Often agencies use these judgments as a basis for determining how to apply the law in the future. These precedents are watched for guidance in future procurements.

Government versus Private Contracting

Contracts have been around for hundreds of years. In simplest terms, a contract is an agreement between two or more parties that can be enforced by a government body. Under the current legal definition, a contract must have four elements to be legally enforced. First, the contract must be between two or more legal entities, individuals or businesses. A person cannot have a contract with himself/herself. Second, all parties must be able to legally enter into the contract and do so freely. The person cannot be coerced, underage, or incapacitated, etc. Third, the contract must have consideration. Consideration is an exchange of something of value. Most courts uphold that it should be an equitable exchange of value to be legal. For example, a contract to purchase a house for a dollar probably would not be enforced in court because the house would be worth more than a dollar, regardless of the condition. As a result, this would not be considered an equitable exchange of value. The most common consideration is exchange of a good or service for a payment of some type. Finally, the substance of the contract must be legal. The court will not uphold a contract for someone to have to perform an illegal act. In private contracts, anything else is allowed. Given these elements, a contract is formed when one party makes an offer and the other party accepts the offer.

As long as the individuals do not break any laws, other salesmanship or common practices are allowed. A common one that is allowed in private contracting and not allowed in government contracting is the “wining and dining” of prospective clients. Another element that is in government contracting that is not present in private contracts is the need for fair competition. In the private sector one company can choose to contract with another simply because the owner likes the other company. This environment is not present in government contracting. The primary reason is the government operated under different rules than the private sector when it comes to contracting. Just as the cricket game mentioned earlier, at first glance it looks very similar to the “same old game” we all know. However, this game has different rules that can be complicated to a new “player.” This complex game can influence some players to stay off the field. Unfortunately, the government can lose out when companies do not want to enter contracting due to this “red tape.” So, if contracting can be simple between private companies, why is government contracting so complicated?



Sources of Government Acquisition Rules

While the “three branches system” of Government helps protect the people from abuses of power with its “checks and balances,” the system complicates the “rule of law” under which government acquisition professionals operate. Since the “checks and balances” gives each branch some influence over the other branches, procurement “rules” are influenced by all three branches. Procurement rules can originate in any of the three branches of government.

The executive branch issues policies and orders that must be followed by all the agencies within the branch. Similar to how a CEO directs the company on how it will conduct business, the Office of the President directs the Agencies on how to conduct different aspects of running the government. In relation to contracting, most of the time these policies are issued from the Office of Federal Procurement Policy (OFPP) under the President. The regulations for procurement are developed in this branch.

With a new President every four to eight years, executive policies change. Just as each CEO has different management philosophies, each President has his own management policy. The management policy and acquisition philosophy of the President affects the policies coming out to the Administration. The larger the difference between philosophies of different Presidents, the greater the policy changes will be when a new President takes control.

The individual agencies issues guidance on how procurement is conducted within the agency. Each agency has a specific mission. Their specific missions influence the needs of the agency. As a result, some agencies procure goods and services that other agencies do not procure. In addition, different agencies operate in different environments. Each agency must have its own guidance on procurement issue. The individual agencies accomplish this guidance in a few ways. Most agencies issue a supplement to the *Federal Acquisition Regulation (FAR)*. In addition, they issue internal Informational Letters and handbooks or directives. These methods allow the agencies flexibility in providing internal guidance for procurement issues. However, agencies have to watch executive and legislative directions. The agencies operate in a dual reporting system. As part of the executive branch, the agencies report to the President. In addition, the heads of the agencies are often called to report or testify to Congress. This puts a unique responsibility on the agencies in developing its internal policies.

The Legislative Branch affects government acquisitions a couple ways. First, this branch writes and passes the laws. All citizens and legal entities must follow the laws passed by the government to the degree it pertains to them. Due to the protections the Founders put in place with the interactions of government and citizens, the Federal Government can put restrictions and laws in place that affect how it conducts business easier than laws that affect private companies and citizens. When a law is put in place that directs the government to conduct business in a certain manner, the agencies must follow the laws that affect its contracting. In general, when society decides it needs to change a rule because of its impact on the people, it creates a law to restrict a particular behavior. The government is the largest buyer of goods and services in the marketplace. With this status comes a lot of power in negotiations. Since the government is concerned with the treatment of the citizens, a number of procurement laws over the years have focused on making the contracting process fair for the companies competing for the government’s business.



Another way the Legislative Branch affects government acquisitions is through control of the money for acquisitions. The Constitution states that no money will be spent by the government that is not appropriated by Congress. The appropriations process is accomplished each year through the Federal Budget. This budget authorizes how much each agency can spend. The agencies then decided to split its money up among its program offices. A Contracting Officer is not allowed to enter into a contract for which the money is not available to be obligated. During the budget process, the Congress will often set certain monies aside for a specific purpose. This practice is called earmarks. The money is released to the agency, but the accompanying law says the agency must use it for a specific purpose.

Laws are often written and passed due to current events in society. Unfortunately, other influences affect the creation of laws. Special interest groups push the agenda of their perspective interests. Special interest groups are not just large organizations. Any group with a specific mission is a special interest group. Veteran groups, environmental groups, small business advocates, etc. are all considered special interests. Some interest groups have more influence than others do. These groups strongly encourage Congress to pass laws that assists their mission. Of course, these special interest groups compete with other groups when it comes to influencing Congress. This ebb and flow of power as laws are passed create new procurement policies that change the rules of how items are purchased. These changes can seem to contradict the rules in place prior to the new law. Another influence on the laws that Congress passes is the “TV effect.” The TV effect is change in public opinion after a “news” story comes out. Previous news stories of problems in the acquisition community have led to Congress investigating and often creating a new law on how procurement is accomplished.

Unfortunately, since laws have a political influence in its creation, flaws in the system exist. In response to current concerns of their constituents, Congressional politicians bring forward new laws. When new laws are presented, the effect on current laws are not always addressed. At times, this process creates conflicting guidance. If a new law is passed, the government must follow the new law. However, current laws stay in effect until they are changed or repealed. So often, a new law that is passed will mandate an action that is contrary to an existing law. This is where the confusion originates for the Contracting Officer in the field.

The language of laws adds to the difficulty in interpreting the intent and how the Contracting Officer is supposed to implement the new law. The nature of laws is to write the law in sufficiently vague language. This process is done on purpose. The reason for this is so the law can be applied to different situations. If the language is too specific, it cannot be applied to unforeseen circumstances. A side effect of this process is that laws must be interpreted. The interpretation of laws is a primary responsibility of the Judiciary Branch.

The Judicial Branch also influences the acquisition process. Most of the states in the US follow common law. Common law is derived from the practices and court rulings. These rulings create a precedence that other courts take into account in future cases. Since these rulings can influence future decisions a court makes, the acquisition community closely follows the court decisions on topics related to contract law. Two types of decisions from the courts affect acquisitions. The first is if a law is challenged as unconstitutional, it could be heard by the US Supreme Court. If any law is determined to be unconstitutional, the law is overturned and people no longer have to follow that law. This is true of any law, including procurement laws. The other way decisions affect procurement is when a claim of unfair treatment is filed by a party harmed by Government in the award of a contract to another party.



Protests and Disputes

Each branch has its processes for hearing complaints that a vendor was treated unfairly in competition for a contract. In the Executive Branch, the vendor files an agency level protest. This protest goes to the Contracting Officer or one level above. If the agency determines the protest is valid, it can overturn the Contracting Officer's decision. In the Legislative Branch, protests are filed with the Government Accountability Office, or GAO. The GAO is a nonpartisan group that reports to Congress. The GAO reports its findings and recommendations to the agency, whether it upholds a protest or not. While the agency does not have to implement the GAO recommendations, Congress watches how much agencies follow these recommendations. The Judiciary Branch hears complaints through the court system. Currently, the court structure that claims progress through is the Civilian Board of Contract Appeals (CBCA). This court was established by section 847 of the *National Defense Authorization Act for 2006*. This is a continuation of the authority to hear disputes between the individuals and the executive agencies established under the *Contract Disputes Act of 1978* (www.cbca.gsa.gov).

The court makes a ruling on the claim in favor of the company or for the Government. If the contractor wins, the court will normally award damages. In these decisions, the court normally states a legal opinion or rationale for its decision. The procurement community pays attention to these decisions. The result of these decisions from the different branches can give insight into what was intended by the law or policy that was the topic of the protest or claim. The legal opinion or decision can give great insight into how future decisions will be made. If a Contracting Officer encounters the same policy or law in the procurement of a good or service, he or she will be able to address the situation properly.

History of Procurement Policy

Since the founding of the United States, the government has paid for goods and services from individuals. As stewards of the people's money, the government must ensure it spends the money wisely. Over time, the government has become the largest buyer of goods and services in the country. With that status, comes a power that could be abused. As a result, the government has passed laws to protect the citizens' interests. These laws protect those doing business with the government and ensure the government is getting a value for the people's money. As problems have developed, actual or perceived, policies have been changed. Contracting policies can be divided into three main historical periods, pre-WWI, WWII, and current.

Prior to WWI, the government started enacting laws to guide its contracting activity with the public. Most of the needs in this time dealt with defense. Some major enactments (PWC, 2008) include *Advertising and Sealed Bids*²⁷ in 1842, *Civil Sundry*²⁸ in 1861, and the *Antifraud Act*²⁹ of 1862. From 1862 to 1921, not a lot of activity occurred in procurement policy. In the pre- and post-WWII period, some additional issues became known. As a result,

²⁷ Required advertising, sealed bids, formalized procedures, and regulations to ensure accountability and reasonable prices (DAU, 2008).

²⁸ Set up a competitive open bid system and formal advertising.

²⁹ Requires all contracts to be written, signed and kept on file to help eliminate carelessness and corruption.



several policies were enacted. In 1921 the *Economy Act*³⁰ was passed. The *Davis-Bacon Act*,³¹ *Buy American Act*,³² and *Miller Act*³³ were passed in 1931, 1933, and 1935, respectively. As the war moved to an end, the *War Powers Act*³⁴ and *Berry Amendments*³⁵ were enacted in 1941. After the war in 1947, the *Armed Services Act*³⁶ was passed in 1947.

The 1970 Commission on Government Procurement³⁷ starts recent history with a study of how the government procures goods and services and how to improve the process. In 1984, the *Competition in Contracting Act* was passed and the *Federal Acquisition Regulation (FAR)* was developed. These two policy changes resulted in a mandate for all contractors to be treated fairly and on a level playing field. In addition, the *FAR* set a policy for all government agencies to follow in procurement of goods and services. In 1996, the Packard Commission³⁸ studied successful companies to determine what the Government could learn from industry. This starts a shift in government to incorporating best practices into how it manages the large agencies. *The National Performance Review*³⁹ of 1993, the *Federal Acquisition Streamlining Act*⁴⁰ of 1994, and *Federal Acquisition Reform Act*⁴¹ of 1996 start addressing the issues companies have in doing business with the government. They also bring to light why companies are able to keep up with changing industries better than government. The processes and bureaucracy was adding a lot of waste to the process, especially with commercial items. A number of the requirements present in a noncommercial item are unnecessary when acquiring a commercial item that has been tested and available to the public.

More recently, the *Defense Transformation Act of 2004* and *Service Acquisition Reform Act Final Report*⁴² of 2007 have influenced government contracting (Vincent, 2009, Slide 8). The *American Recovery and Reinvestment Act of 2009* added a new emphasis on

³⁰ Established the ability of an agency to purchase products or services from another agency.

³¹ Requires contractors to pay the “prevailing rate” to workers in construction.

³² Sets a preference for American products. It was established during the Great Depression to protect American companies.

³³ Set requirements for payment and performance bonds to protect government interests under construction contracts over \$100,000.

³⁴ Removed most statutory requirements to allow rapid mobilization of industry during WWII. The end result was a more distant relationship between government and contractors.

³⁵ Required the Department of War (now the DoD) to buy domestic end products.

³⁶ Established extensive regulations for military procurements. Gave the DoD authority to contract for National Defense needs.

³⁷ After WWII, increase in regulations. Took away much of the discretion of the Contracting Officer that was granted under the *War Powers Act*.

³⁸ Determined many problems in procurement were a result of too much regulation over the decades.

³⁹ Focused on reform and reducing costs.

⁴⁰ Focus on efficiency in contracting. Eliminates some regulations and restrictions when procuring commercial items and set up simplified procedures.

⁴¹ Established dispute resolution procedures, setting competitive ranges and allowed rejected offerors to request debriefings.

⁴² Revised definition of commercial services and incorporated best practices from industry into the government procurement process.



transparency and reporting. The implementation of the reporting aspect for vendors has proven challenging to companies, especially one new to government contracting.

Conflict in Guidance

This concern for fair treatment is present in a number of the laws affecting the Federal Procurement process. Such laws as the *Competition in Contracting Act*⁴³ (CICA) and *Truth in Negotiation Act*⁴⁴ (TINA) address how the Government interacts with prospective vendors. The *Small Business Act* put an emphasis on contracting with small disadvantaged business. However, the Government has a fiduciary responsibility to spend the taxpayers' money wisely. The Government has had to balance using economies of scale to secure a better price for the taxpayers and assisting socioeconomic groups. As issues become a higher priority for the Government, it will get more attention from the President and rules or laws are passed to focus more of the contracting business on that area. A recent example of this is "green," or energy efficient, products⁴⁵. Guidance has come out recently on green contracting and the use of more Energy Efficient equipment. However, the *FAR* directs the Government to avoid using Brand Names as much as possible. Since "green" technology is new, this emphasis can limit competition to specific brands.

Another priority for the Government is the veterans returning from war zones. Recently Public Law 109-461 includes the Veteran's First Program for the Department of Veteran Affairs (VA). This allows the VA to choose Service Disabled Veteran Owned Small Business (SDVOSB) or Veteran Owned Small Business (VOSB) before any other groups.⁴⁶ The conflict arises when the SDVOSB does not sell the more energy efficient product. Then the Government Official needs to make a decision between the SDVOSB and the more energy efficient product that may be only made by a large business. Another conflict occurs when a small business gets a contract due to competition restrictions, but does not add value to the final product or service. The government must balance the need for spending money wisely and its social responsibility to help small businesses. In addition, the priority order under the socioeconomic programs changes over time. These changes result from the small business groups fighting, through special interest groups, for a better position in the order of priority.

Recently, the government has shifted to more of a focus on using fixed price contracting whenever possible. This is supposed to lower risk and expense for the Government. However, by shifting risk onto the contractor, the vendor considers that risk when pricing their product or service. With commercial items, this makes more sense, because a company can better gauge its cost due to the large volume of an item or service it has already sold. Similarly, once a noncommercial product is developed for the Government and had several production runs, the contractor should have a better control on its cost and the Government can procure future production on a fixed price basis.

With services, like construction, the ability to fix the price becomes a little more risky. Even with a set service, the conditions of every project are different. In essence, the

⁴³ Increase competition to more companies resulting in lower prices for the government.

⁴⁴ Requires contractors to submit cost or pricing data and certify data is current, accurate and complete.

⁴⁵ Energy Star programs require items on this list to meet certain energy efficiency standards.

⁴⁶ This reorganizing of priority groups is specific to the VA and as a statute overrules the *FAR* guidance.



“commercial product” is being tailored. In addition, due to the *Service Contract Act*, a firm fixed price service contract is not as firm as a commercial item. When wage determinations change in an area a service is provided, the contractor must adjust wages. This is the one time a contractor can make a change in which the Government really does not have a choice. When the contractor requests an equitable adjustment due to change in wage determination, the government needs to adjust the pricing of the contract, due to law changes. The contract is a firm fixed price in name early. In practice, the contract acts like a fixed price with economic adjustment. The adjustment is tied to the wage determination. The current contracting environment created by these varied laws and regulations have created a jungle for the Contracting Officer to navigate (See Figure 1).

The ideal place to address these conflicting requirements is in the legislative and policymaking level. At the policy level, the *FAR Council*⁴⁷ issues proposed rules to the public and allows everyone to comment on the policy before making a final rule. This process helps to bring to light possible conflicts that the new policy could create. However, policies must work within the legislation. If the two are at odds, the law trumps policy. In an ideal process, a study would be done before enacting new laws affecting procurement. A similar period for public comment before the final law is drafted would put a spotlight on some of these issues before the government encounters them in practice. This process would greatly improve the procurement law and policy processes. This new initiative is contrary to how laws are normally made, with lawmakers comprising on different aspects and combining bills to accomplish their goals. However, this new approach would fit in with the new focus on government transparency and use of technology.

Contract administration is another area that has conflicting guidance. Contracts must be properly administered. This is especially true in service or long term contracts. The success of the contract is how the service is delivered. However, many agencies put a focus on activity up to award. Many Contracting Officers are evaluated on contracts awarded, and time to get to award. Direction is given to improve administration, but no resources are expanded to assist with contract administration.

A final example of a conflict in guidance is the directive from the Administration early last year that new service contractors must give “first hire” option to the employees of the incumbent contractor. The purpose of this directive is to minimize the impact of changing vendors, by having individuals familiar with the services and location stay in place to provide the services to the Government. However, this can conflict with the Contracting Officer’s fiduciary responsibility to spend the taxpayers’ money wisely. If a current contractor is not performing well, then the Government can be stuck with the same underperforming employees under a new contractor. The government made a best value decision to select a company, when it is receiving the employees of a different company, if they accept the offers of the new vendor.

Approaches

A few approaches are currently used to address conflicting requirements. As laws are passed or directives are received by the President, the Office of Federal Procurement Policy starts the process of examining how to implement the new directive or law into the current Federal Policies. Initially, directives are sent to agencies to start implementing a policy while a proposed change is debated. Once a final rule is made, it is incorporated into the Federal Policy. This process is why different parts of the *FAR* seem to give different

⁴⁷ The governing body that issues changes to the *Federal Acquisition Regulation*.



guidance from other parts. The result is like several people each writing a different chapter of a book and then putting it together. Periodically, a full review of the *FAR* in whole and how directions affect other parts needs to be accomplished.

One way to address different competition requirements has been the use of thresholds. Under the micro-purchase threshold, the customer can use a government credit card to procure the goods or services needed. Between the micro and simplified thresholds, all procurements should be reserved for small businesses to the greatest extent possible. Above the simplified threshold, the agency should set aside for small businesses in specific socioeconomic groups whenever possible. If the agency must go to a large business, the large business should subcontract some of the work to a small business. A final approach used is writing an exception into the regulation (i.e., exception to the non-manufacture rule). These approaches help to satisfy a specific requirement, but not another. For example, under the *Competition in Contracting Act* all companies should have equal opportunity to bid on contracts. However, the focus on small businesses shuts large businesses out of competition in a number of contracts. If thresholds are used to determine set aside, the spirit of competition in contracting should require procurements above a certain amount be automatically open to large business.

Currently the decision on whether small businesses can compete or the competition should be opened up to larger businesses is left up to the agency through the Contracting Officer and the Review Boards within the agency. As Contracting Officers become better skilled and trained, they can become the trusted business advisors that are needed to make those types of decisions.

Another way to improve efficiency in the contracting field is to focus on training and accountability at the Contracting Officer level. Many agencies have internal policies of higher approval for contracting actions. This oversight is intended to ensure the Contracting Officer's compliance with policies. Compliance with policy is only a problem if one of two causes is present. Either, the contracting officer displays a laxness in the responsibilities of the job; or the person lacks the knowledge and experience to implement the policy properly in the procurement process. The approach best suited to solving this issue is to focus on training and accountability. If organizations ensure the proper, continuing education and hold the Contracting Officer accountable for the actions, a lot of the bureaucracy can be eliminated.

Not all the fixes have to be done above the Contracting Officer. The *FAR* Guiding Principles state "unless a strategy is expressly forbidden by statute or regulation, the strategy is open to the Contracting Officer." This gives the Contracting Officer a lot of leeway in how to conduct business. The Contracting Officer can use some innovative techniques to satisfy these different requirements and provide the best value for the Government. One method to address recurring needs has been establishing an IDIQ.⁴⁸ The primary advantage to an IDIQ is the Government does an evaluation once for multiple requiring requirements anticipated in the near future. This streamlines the process of future procurements because the awardees have already been evaluated for past performance and technical ability to fulfill the requirements. When a future request from a customer occurs, an order can be placed without having to compete a second time. Eliminating this need for future competitions can speed up the procurement process without sacrificing quality. While this approach works well for most commercial items, problems occur with large quantities and small businesses or with services. Since services are customized, the final price could be

⁴⁸ Indefinite Delivery Indefinite Quantity contract

different from one order to another. This removes the advantage of streamlining the procurement, the primary advantage of an IDIQ. One product recently used in construction that can be easily used in services is a MATOC⁴⁹. This contract type issues awards to multiple vendors for an expected series of requirements. The same initial evaluation of an IDIQ occurs with a MATOC. The difference between the two is a pool of prequalified contractors is available under a MATOC. The MATOC uses either a seed, or first task order, project or a fictional project within the original solicitation to allow for a proper technical evaluation. This allows the Government to evaluate approach on customized requirements, such as construction or services. The way to use this and still keep competition is to hold competition at the task order level among all the awardees of the MATOC. This ensures the government receives the benefits of a small group of prequalified vendors bidding against each other for the service.

The use of options⁵⁰ in a contract has become more popular for ongoing services. This approach works best in services that are required from one year to the next and do not change rapidly in price and specifications. In areas where prices are rapidly dropping, long options are not in the best interest of the Government. If prices are rising rapidly, the contractor is put at a disadvantage. Technology is a prime area where specifications are constantly changing. An IDIQ for installing equipment might be better than option years in this case.

One very effective technique to addressing requirements, when done properly is performance based contracting. This puts a focus on the desired outcome, not how it is accomplished. This lesson was learned from private industry. Companies often present a problem to a potential vendor as “this is what I want it to do, I don’t care how it does it.” Success is measured by the result. The Government took a while to arrive at this same conclusion, but is now taking greater interest in this approach.

In contract administration, focus has shifted back and forth between “cradle to grave” and a separate Administering Contracting Officer (ACO). Both have advantages. The problem with “cradle to grave” is the focus on Contracting Officers to concentrate on awarding new contracts and the original CO may no longer be with the organization. Since most employee evaluation is directed at this phase, administration gets less attention. However, in the ACO structure, often the ACO did not write the contract and may be in a different location. One way to address this is to adopt the ACO model, but have one at each contracting shop at the base level. This way a focus is on contract administration and the ACO is accessible at the location. In addition, if any questions arise the ACO can ask the CO that wrote the contract.

So how can the Government capitalize on economies of scale and still nurture small businesses. Strategic Sourcing⁵¹ has been offered as a way to leverage the buying power of the Government. However, this eliminates most small businesses from competing. A couple methods can be used to accomplish both benefits. First, procurements under strategic

⁴⁹ Multiple Award Task Order Contract

⁵⁰ Ability to extend a contract for a period of performance or quantity that was agreed to in the original contract.

⁵¹ A directive that agencies put in place contracting vehicles to procure large quantities of products continually bought across the Government, such as office supplies and computers, etc.

sourcing can encourage teaming⁵² among companies to provide the products or services. Under this approach, companies can band together in larger joint ventures to handle the large volumes. Another approach is for the agencies to set up large, regional Multiple Award IDIQs to handle these requirements. This approach would set up a “mini-schedule” similar to GSA. While these currently exist as GWACs⁵³, the communication of availability to contracting offices is lacking. Better use of these contracting instruments can balance the requirements.

Better reporting and contract administration can help ensure the government receives good services. The Presidential directive that new contractors give “first hire” choice to the incumbent’s employees can cause some quality issues. While the intent is to minimize delivery of services by having people on the ground that already know the job, often performance issues are rooted in the employee performing the service not necessarily the management. The Government recently expanded the past performance database by requiring all agencies to utilize PIPRS. Expanding this system to allow the Government to report the performance of individuals would help to ensure the government does not inherit the same bad employees under a new company. In order for this to work, the Government must report bad performance when it occurs. Too often the Government would prefer to let the contract end or Terminate for Convenience because it is the route of least resistance. The Government does a disservice to the taxpayer when neglecting this responsibility.

Conclusion

Contracting Officers operate in a complex environment. Juggling requirements becomes the real art form of Government contracting. At the same time, agencies want to see new innovative methods to purchasing goods and services that will save the taxpayer money. Several methods are available to the well-informed contracting officer to fulfill the socioeconomic goals and be innovative and efficient. One example of this is the MATOC contract for construction. With some creativity, this type of contract can be used for other activities, like services. By using innovative processes; an agency can leverage its buying power over the year. In addition, the agency could make this a set-aside for a specific socioeconomic group if market research shows enough competition. As contracting policies change over time, the Contracting Officer must adapt operations to satisfy customer needs and stay within the regulatory guidelines.

References

- CBCA. (2009, December). *United States civilian board of contract appeals website*. Retrieved March 15, 2010, from <http://www.cbca.gsa.gov>
- DAU. (2008, October). History of American acquisition (Coursework, Class CON 100, Lesson 1). Frederick, MD: Veteran Affairs Acquisition Academy.
- PWC. (2008, October). VAAA intern development program (Coursework, SBLD 008-A). Frederick, MD: Veteran Affairs Acquisition Academy.
- Vincent, L. (2009, December 29). Government contracting seminar slides. Slide presentation.

⁵² An arrangement where more than one company cooperates on a project. In a number of teaming agreements, the Government has a direct contractual relationship with the members of the team, unlike a prime/sub relationship.

⁵³ Government Wide Acquisition Contract



Contracting Environment

Oversight and Guidance

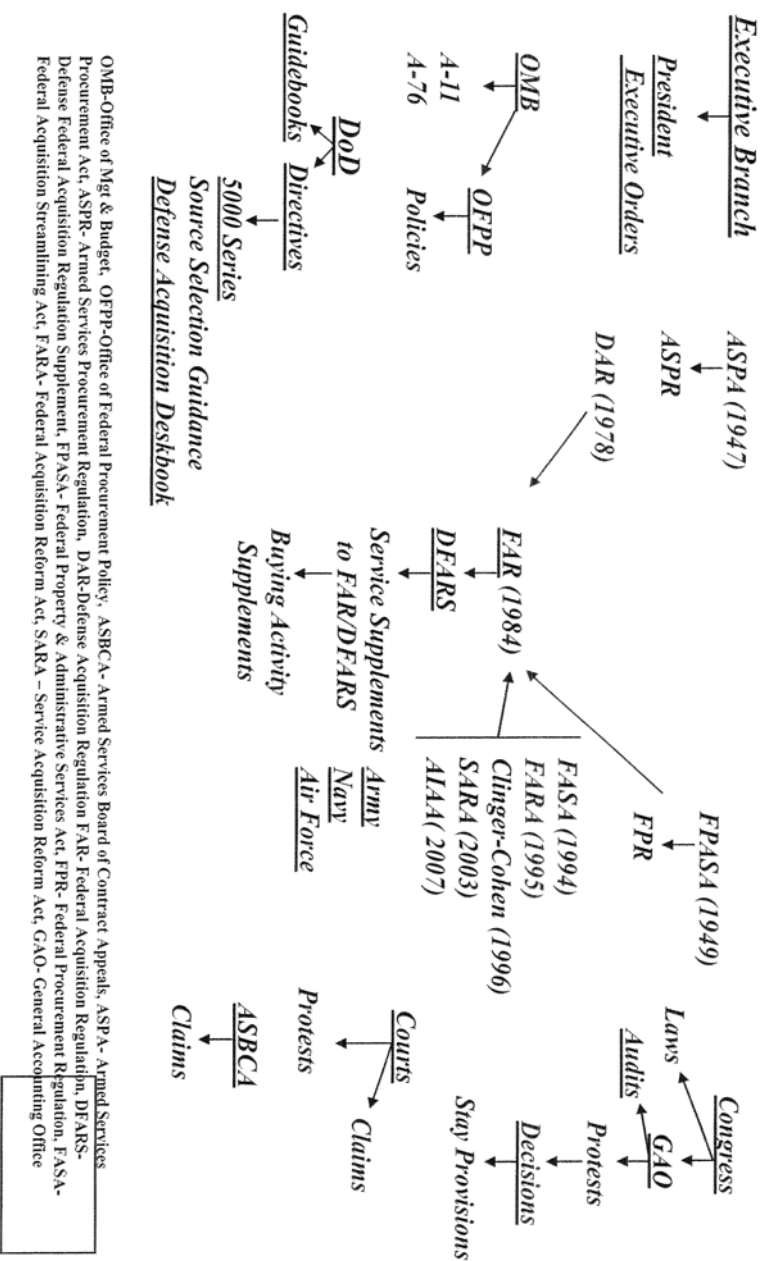


Figure 1. Contracting Environment (Vincent, 2009)

Services Supply Chain in the Department of Defense: Comparison and Analysis of Acquisition Management in the Army, Navy, and Air Force

Aruna Apte—Dr. Aruna Apte is an Assistant Professor in the Operations and Logistics Management Department, Graduate School of Business and Public Policy, at the Naval Postgraduate School, Monterey, CA. Her research interests are in the areas of developing mathematical models and algorithms for complex, real-world operational problems using techniques of optimization. It is important to her that her research is directly applicable to practical problems and has significant value-adding potential. She has numerous publications in peer-reviewed journals. She teaches a mathematical modeling course and has advised over 30 students for theses and MBA reports. Currently, she is working in the areas of developing mathematical programming models in humanitarian logistics and military logistics. Before NPS, she worked as a consultant at MCI and taught at Southern Methodist University. For more information, please visit http://research.nps.edu/cgi-bin/vita.cgi?p=display_vita&id=1105652618

Uday M. Apte—Dr. Uday M. Apte is Professor of Operations Management at the Graduate School of Business and Public Policy, Naval Postgraduate School, Monterey, CA. Before joining NPS, Dr. Apte taught at the Wharton School, University of Pennsylvania, PA, and at the Cox School of Business, Southern Methodist University, Dallas, TX. Dr. Apte holds a PhD in Decision Sciences from The Wharton School, University of Pennsylvania. Prior to his career in academia, Dr. Apte worked for over ten years in managing operations and information systems in the financial services and utility industries. Since then, he has consulted with several major US corporations and international organizations. Dr. Apte has served as a founder and President of the College of Service Operations, Production and Operations Management Society (POMS), and as a board member and vice president of POMS. Areas of Dr. Apte's research interests include managing service operations, supply chain management, and globalization of information-intensive services. He has published two books and over 40 articles, five of which have won awards from professional societies.

Rene G. Rendon—Dr. Rene G. Rendon is an Associate Professor at the Naval Postgraduate School where he teaches defense acquisition courses. Prior to his appointment at the NPS, he served for more than 22 years as an acquisition and contracting officer in the United States Air Force, retiring at the rank of lieutenant colonel. His Air Force career included assignments as a contracting officer for the Peacekeeper ICBM, Maverick Missile, and the F-22 Raptor. He was also the director of contracting for the Air Force's Space Based Infrared satellite program, and the Evolved Expendable Launch Vehicle rocket program. Rendon's publications include *Management of Defense Acquisition Projects* (2008), *Government Contracting Basics* (2007), *US Military Program Management: Lessons Learned & Best Practices* (2007), and *Contract Management Organizational Assessment Tools* (2005). He has also published in the *Journal of Public Procurement*, the *Journal of Contract Management*, and the *Project Management Journal*.

Abstract

This paper presents the results of our empirical studies of current management practices in services acquisition in the Army, Navy, and Air Force. The primary objective of these studies was to develop a comprehensive understanding of how services acquisition is being managed within, as well as across, individual military services. In these empirical studies, we developed and deployed a Web-based survey to collect primary data. Specifically, we studied the current management practices in such areas as contract characteristics, and acquisition management methods including regional- or installation-level acquisition, use of project management approach, acquisition leadership and ownership of



requirements. We also studied other program management issues such as scope and ability of personnel responsible for acquisition, adequacy of acquisition billets and their fill rates, and training provided to services acquisition personnel.

We found that for the most part the services contracts awarded and administered conformed to our expectation. For example, most service contracts are competitively bid, fixed-priced awards with a minimal use of any type of contract incentives. The survey data also confirmed that the Navy uses regional approach in services acquisition, while the Army and the Air Force use installation-level approach. These differences, in turn, appear to be having important implications for other acquisition management practices such as the use of project management and contract surveillance. One surprising finding of the study was that the project teams are often led by the contracting officer as opposed to a formally designated project manager responsible for the overall service project success. Finally, the survey respondents indicated that the number of authorized staff positions for services acquisition was inadequate and, furthermore, that the existing billets were inadequately filled.

The analysis and comparison of management practices in different military services was used as the basis to develop, and report in this paper, our preliminary recommendations for improving the management of services supply chain in the Department of Defense.

Keywords: Service Supply Chain, Services Acquisition, Service Lifecycle, Contract Management, Project Management, Program Management

Introduction

The service sector represents the largest and the fastest-growing segment of the economies of the US and other developed countries. For example, in the US, services accounted for roughly 80% of employment in the year 2004 (US Department, 2005). The growth of services in the overall economy is also mirrored by growth of services acquisition in private sector companies (Smeltzer & Ogden, 2002) and in the government. For example, as seen in Figure 1, the procurement of services in the DoD has continued to increase in scope and dollars in the past decade. Even considering the high value of weapon systems and military equipment purchased in recent years, the DoD has spent more on services than on supplies, equipment and systems together (Camm, Blickstein & Venzor, 2004). Specifically, the DoD obligations on contracts have more than doubled between fiscal years 2001 and 2008—to over \$387 billion, with over \$200 billion spent just for services (GAO, 2009c). The procured services presently cover a very broad set of service activities—including information technology and telecommunications services; maintenance and repair of equipment; professional, administrative, and management support; and transportation, travel, and relocation services.



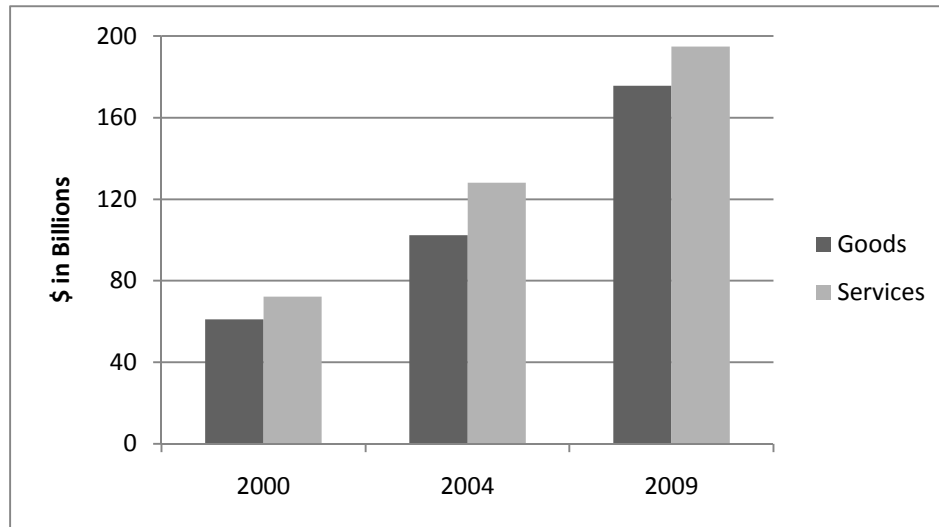


Figure 1. The DoD's Contracts for Goods and Service (2000–2009)

As the DoD's services procurement continues to increase in scope and dollars, the DoD must give greater attention to the management of services contracts. However, the increase in service contracting has coincided with the reduction in the federal government workforce. For example, the size of the federal workforce fell from 2.25 million in 1990 to 1.78 million in 2000 (GAO, 2001). This mismatch between the increasing workload and the decreasing size of the workforce, and the unique nature and complexities associated with services acquisition have possibly created an environment in which following the best practices has not always been feasible. For example, between 2001 and 2009, the Government Accountability Office (GAO) has issued 16 reports related to trends, challenges, and deficiencies in contracting for services. In addition, between 2002 and 2008, the DoD Inspector General (DoD IG) has issued 142 reports on deficiencies noted in the DoD acquisition and contract administration process. Both the GAO and DoD IG have identified market research, contract type, project management, requirements management, personnel training, and contractor oversight as just some of the critical deficient areas in services contracts. Further discussion of these deficiencies is provided in Exhibit 1.

- The government is required to conduct *market research* to determine the market's capability for providing the required supply or service and the government's appropriate contracting strategy for the procurement (Rendon & Snider, 2008). Reports have shown that the DoD has not conducted adequate market research during procurement planning of services contracts (GAO, 2002a; DoD IG, 2009).
- Selecting the appropriate *contract type* is essential for ensuring the appropriate sharing and allocation of risk between the government and the contractor. Fixed-price contracts allocate the majority of the cost risk to the contractor, while cost reimbursement contracts provide for most of the cost risk to be borne by the government. Government reports have shown that inappropriate contract types were used in services contracts, resulting in more risk to the government (GAO, 2001; DoD IG, 2009).
- The use of *project management* tools and techniques, such as designated formal project managers, project teams, and project lifecycles, have been considered a best practice in managing service contracts. GAO reports have shown that the DoD lacks the proper management structure and processes for managing services contracts (GAO, 2007b; DoD IG, 2009).
- Sufficient *requirements management* is essential for identification and development of needs for the DoD, in terms of required services. If requirement management is insufficient, the resulting service contracts will not adequately meet the customer's needs. The GAO and DoD IG reports have identified poorly defined requirements and insufficient requirements management as problems in service contracts (GAO, 2007b; DoD-IG, 2009).
- Defense *contract management* requires specialized skills and competencies that come from extensive *training and experience*. A properly trained and competent acquisition workforce is considered the heart of successful defense acquisition management. With the downsizing of the DoD workforce, the lack of a qualified acquisition and contracting workforce to manage the increased workload in DoD service contracts continues to plague DoD service contracting efforts (GAO, 2002b; GAO, 2009b).
- The essence of DoD *contract management* is the proper administration of contracts and *oversight of contractor performance*. The lack of effective contract administration and contractor oversight increases the government risk of not ensuring total value for the dollars spent on service contracts. GAO and DoD IG reports have consistently identified contract administration and contractor oversight as problem areas in the management of services contracts (GAO, 2005; GAO, 2007a; GAO, 2007b; DoD IG, 2009).

Exhibit 1. Deficiencies in Services Contracting

Indeed, DoD contract management has been listed as a “high-risk” area by the GAO since 1992 (GAO, 2009a). This “high-risk” status reflects the DoD’s challenges in achieving their desired outcomes in terms of meeting the service procurement cost, schedule, and performance objectives. The DoD is at risk of paying higher prices for services than necessary. Recently, the DoD Director of Defense Procurement and Acquisition Policy (DPAP) identified the inappropriate use of service contracts in the DoD (Director, 2007) and is planning to take actions to improve contracting for services throughout the DoD (Director, 2006).

Service production differs from manufacturing of products in several ways due to distinguishing characteristics of services. There is a growing body of literature on operations management in service firms. The key characteristics of services discussed in



textbooks (Fitzsimmons & Fitzsimmons, 2006; Metters, King-Metters & Pullman, 2003) include the intangibility of service output, co-production, simultaneity of production and consumption, the inability to store services, and the complexity in the definition and measurement of services. These characteristics also lead to differences in the marketing of services (Lovelock, 1992; Hutt & Spec, 1998).

Given these differences in the production and marketing of services as opposed to that of manufactured products, it is natural to ask if the acquisition of services is essentially the same as acquisition of products or if differences exist. And, if the differences exist, then what they are, in general and for specific services, and what do they imply for the management of services acquisition? Given the growth in size and scope of services acquisition in today's economy, these questions are undoubtedly important.

A survey of academic literature indicates that there exist only a handful of studies aimed at addressing some of these questions. For example, Smeltzer and Ogden (2002) examined purchasing professionals' perceived differences between purchasing materials and purchasing services; Ellram, Tate and Billington (2004) developed a supply chain framework appropriate for a services supply chain by comparing and contrasting the applicability of three product-based manufacturing models; and Schiele and McCue (2006) studied the acquisition of consulting services in public sector. Although these and other studies have started to address some of the questions identified above, for the most part, the important questions raised above remain unanswered. Furthermore, given the peculiarities of government procurement and the GAO and DoD IG reports on the deficiencies in the DoD acquisition and contract administration processes, there exists a unique and significant opportunity for conducting research into the management of services acquisition in the DoD.

We have addressed the need for research in the area of services acquisition by undertaking a series of research projects. The first two research projects were exploratory in nature. In the first project, we tried to understand the major challenges and opportunities in the service supply chain in the DoD (Apte, Ferrer, Lewis & Rendon, 2006) by undertaking in-depth case studies on acquisition of services in three different organizations: Presidio of Monterey, Travis Air Force Base, and the Naval Support Detachment Monterey (NSDM). The second research project was targeted at studying the program management infrastructure (Apte & Rendon, 2007). In this research, too, we conducted two additional in-depth case studies of innovative project management approaches at the Air Education and Training Command (AETC) and at Air Combat Command (ACC).

The next two research projects were survey-based empirical studies aimed at developing a more comprehensive understanding of how services acquisition is being managed at a wide range of Army, Navy and Air Force installations. Specifically, the third research project was aimed at understanding management of services acquisition in the Navy and the Air Force (Apte, Apte & Rendon, 2008), while the fourth research project was aimed at the Army contracting centers (Apte, Apte & Rendon, 2009).

The objective of the fifth research project, the preliminary results of which are being reported in this report, is to analyze the primary data collected in earlier empirical studies involving the Army, Navy and Air Force and to compare the results so as to develop a more thorough and comprehensive understanding of how services acquisition is being managed within, as well as across, individual military services. The analysis of survey results will focus on the following areas: contract characteristics, acquisition management methods, project team approach, services acquisition leadership, and other management issues. The results of this analysis and comparison will be used as the basis to develop preliminary,



department-specific recommendations for improving the management of services supply chain.

The paper is organized in four sections, including the current introductory section. In the next section, we describe the empirical studies we conducted, including the survey research methodology used in the study. The results of survey data analysis and some salient observations are provided in the third section. The findings and conclusions of the study and our recommendations for improving services acquisition and for future research are presented in the fourth section. We wish to point out that the tables summarizing the survey data can be found in two previous technical reports, Apte, Apte, and Rendon (2008; 2009).

Research Methodology and the Empirical Studies

The methodology used in this research consisted of a survey instrument specifically developed to address the research objectives and questions mentioned above. This was a Web-based survey instrument developed using the “Survey Monkey” software. The developed survey was first pilot tested for its validity (Compton & Meinshausen, 2007) and was fine tuned prior to its use in the third and the fourth research projects.

The survey begins with questions focusing on specific demographic data and then asks specific questions related to the approach, method, and procedures used in the acquisition of services for specific categories of services. The categories of services targeted in this research are (1) professional, administrative, and management support; (2) maintenance and repair of equipment; (3) data processing and telecommunications; (4) utilities and housekeeping; and (5) transportation and travel. These categories were selected because collectively, they represent a significant fraction of spending for all the services, and are commonly acquired in the Army, Navy and Air Force.

The survey instrument includes core questions related to the methods and procedures used in the acquisition of services for the above service categories. These core questions focus on the following areas:

Contract Characteristics. The purpose of this category of questions is to gain insight into the dominant procurement methods and contract types used in the acquisition of services. The contract characteristics examined in this section are degree of competition (competitively bid or sole-source), contract type (fixed-price or cost-type), and type of contract incentive (incentive-fee, award-fee, or award-term).

Acquisition Management Methods. The purpose of this broad category of questions is to understand the management methods and approaches used in the acquisition of individual services at each phase of the contract management process. For each of the contract management phases, the survey asks whether the phase was conducted at a regional, installation, or some other organizational level. This core question category also focused on whether a project-team approach was typically used in the acquisition of the respective service category. The questions explore the position of the services acquisition project team leader, such as a program/project manager or contracting officer. The questions also explore information on the owner of the requirement for the service being acquired.

Other Program Management Issues. This last category of core questions is focused on the use of a lifecycle approach, length of assignments for services acquisition management personnel, use of market research techniques, level of staffing in services acquisition management, and level of training of services acquisition management



personnel. These questions use a Likert-type scale to measure the level of agreement or disagreement among the respondents' statements.

The questionnaire described above was used to conduct surveys in all three military services. A summarized description of these survey-based empirical studies is given below.

- Army: The standardized survey was deployed to 81 contracting offices. The survey was distributed across 8 major contracting centers throughout the Army, including 40 Army installations. We received a total of 61 responses to the survey, with a survey response rate of 75%.
- Navy: The data was collected in the Navy survey at the installation level. The data inputs were provided by the Navy Regions in charge of the installations in CONUS. We received inputs from six Regions—covering 66 Navy installations plus Naval Supply (NAVSUP) and Naval Medical Logistics Command (NMLC).
- Air Force: The survey instrument was deployed to 50 Air Force Contracting Squadrons, representing six Air Force major commands. There were 34 responses from the survey, resulting in a 68% response rate. These responses represented all six Air Force major commands.

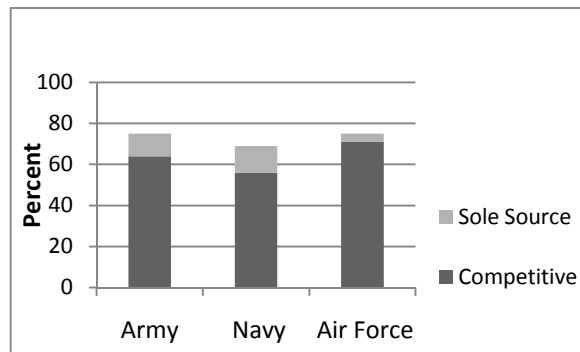
Analysis and Comparison of Survey Data

As mentioned above, a summary of the survey data can be found in the earlier reports related to the third and the fourth research projects. Depending on the nature of responses received during the surveys, the level of details available for individual military services—Army, Navy and Air Force—were somewhat different. Nevertheless, all three surveys generated data with sufficient details so that we were able to compare services acquisition practices across all three services and draw meaningful conclusions. We present below the results of our analysis and comparison of acquisition management practices in the Army, Navy and Air Force through a series of figures arranged along the lines of data categories described in section 2 above.

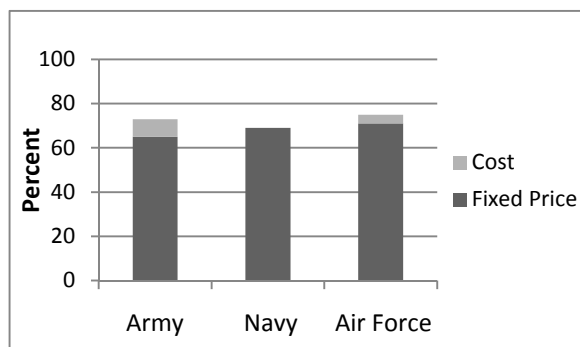
Contract Characteristics

We discuss three aspects of contract characteristics: degree of competition, type of contracts and contract incentives. It should be noted that the Navy and the Air Force surveys were conducted in 2008, while the Army survey was conducted in 2009. Consequently, the Army survey results contain data for 2008, while the data streams for the Navy and the Air Force surveys end in 2007. We used the contract characteristic data for 2007 and computed averages across services and acquisition phases to obtain measures of contract characteristics. The comparison of contract characteristics for Army, Navy and Air Force is depicted in Figure 2.

2A. Degree of Competition



2B. Contract Type



2C. Contract Incentive

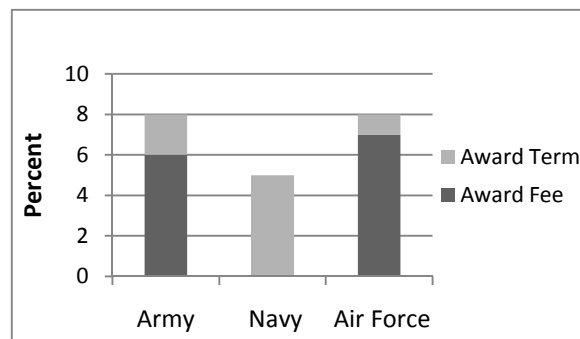


Figure 2. Contract Characteristics

Degree of Competition

Providing for full and open competition is a public policy and statutory requirement in government contracting. The *Competition in Contracting Act of 1984 (CICA)* is a public law, enacted to increase the number of government procurements conducted using the procedures of full and open competition. Unless the government can justify an exception to the competition requirements, the procurement must provide for full and open competition in the solicitation and award of the contract. In addition to supporting accountability and transparency in government contracts, competitive procurements also result in competitively priced proposals that increase the government's ability to negotiate a fair and reasonable contract price.

As we note in Figure 2A, the predominant procurement approach used in services studied was full and open completion. Since these services are traditional and commercial in nature (administrative, maintenance, data processing, utilities/housekeeping, and transportation services), it would follow that the competitive marketplace would be capable of proposing and competing for these contracts. However, we also noted that a small but notable portion of contracts for Navy and Army were sole-sourced. We do not have detailed data on these sole-sourced contracts, but perhaps the services acquired were context-specific and unique in nature.

Contract Type

The *Federal Acquisition Regulation* categorizes the major contract types as fixed-price, and cost-reimbursement. Each of the various types of contracts (specifically cost-reimbursement versus fixed-price contracts) reflects the sharing and allocation of risk between the buyer and the seller. Fixed-priced contracts are appropriate for well-defined requirements in situations with a low risk of performance. Using these types of contracts, the contractor holds the major burden of risk. On the other hand, under cost-reimbursement contracts, which are appropriate for developmental requirements, the risk of performance is high. In these types of contracts, the government shares the major burden of risk. As mentioned above, given the commercial and low-risk nature of services being studied, firm-fixed price contracts would be the appropriate contractual instrument for these service projects. We note in Figure 2B that, as expected, a significant majority of the contracts were fixed-price.

Contract Incentive

In some situations, the government may want to subjectively incentivize the contractor to meet higher levels of performance, which are over and above the basic requirements of the contract. In these situations, award-fee or award-term contract incentives may be used. In award-fee/term contracts, the contractor may earn additional money (or contract periods of performance for service contracts), based on the government's judgmental evaluation of the contractor's superior performance. Since commercial services are usually well understood and the output or outcome can be reasonably well defined, there is less need to use contract incentives. That is what we observe in Figure 2C.

Acquisition Management Methods

In this sub-section we provide comparison of Army, Navy and Air Force practices along two dimensions: the organization level at which services are acquired, and the use of a project team approach. The comparison is shown in Figure 3.



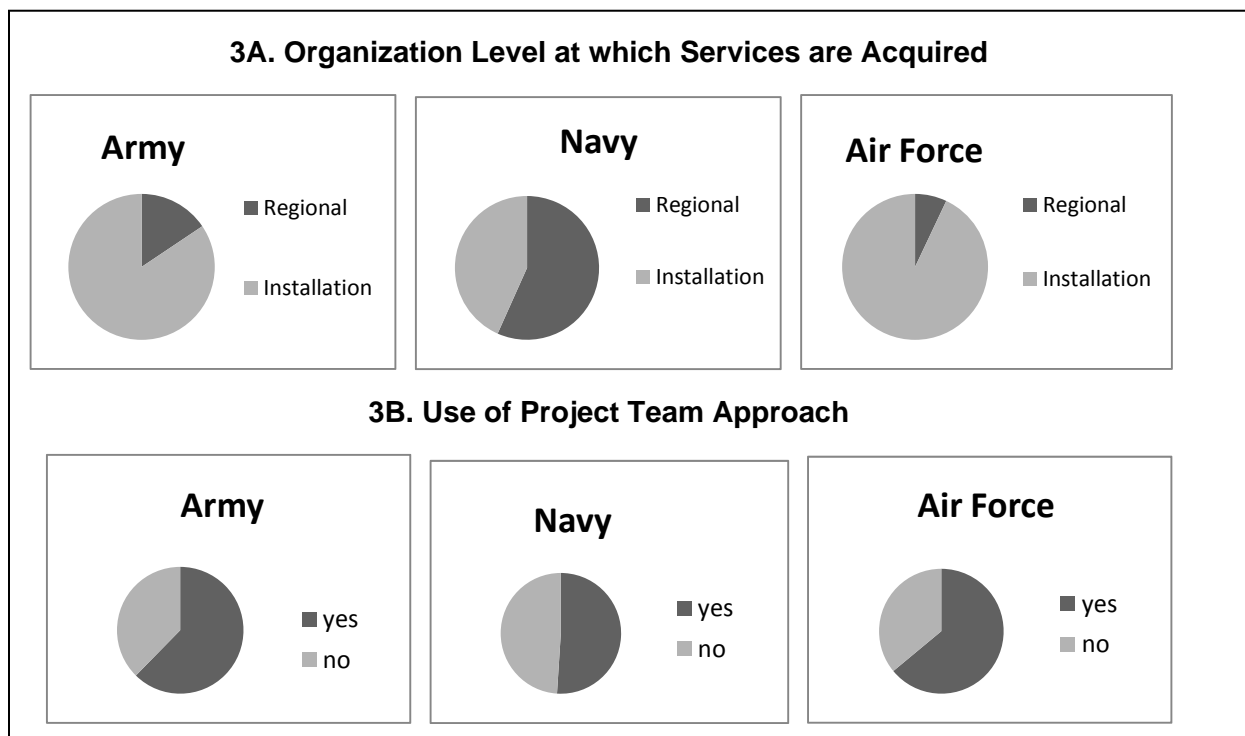


Figure 3. Acquisition Management Methods

Organizational Level at which Services Are Acquired

The military departments procure services and manage service acquisitions at the installation level or regional level. The proximity of where the acquisition and related contracts are managed to where the services are actually performed may have an impact on the effectiveness of the project management, as well as the success of the service project. Services performed at one location, with the contract and overall project managed at a distant location, may result in less than adequate management and control of the project, as well as less than proper surveillance of the service contractor. Insufficient control of the project and less than adequate surveillance of the service contractor increases the risk to the DoD of not receiving the full value of its service procurement dollars.

It is not possible, however, to derive any generalized conclusion that acquiring services using either regional or installation-level approach is necessarily better than the other approach. Regional approach (centralized procurement) can give rise to economies of scale, uniformity of procedures, and possibility of consistently using best acquisition practices. On the other hand, installation-level acquisition (decentralized procurement) allows for easier implementation of project management and program management approaches, including accurate requirements definition and proper surveillance. Under either approach, however, key to success is adopting suitable management practices.

We note in Figure 3A that Navy services acquisition takes place predominantly at the regional level, whereas in the Army and the Air Force, it occurs predominantly at the installation level. This difference in approaches should be kept in mind when analyzing the effectiveness of various management practices discussed later in the section

Project Team Approach

Service acquisitions, such as information technology services or aircraft maintenance services, are typically technically complex and require support from various functional areas, such as engineering, procurement, finance, and logistics. Best practices in project and contract management reflect the use of project teams—specifically cross-functional teams—in the management of service projects. The use of project teams facilitates the proper integration and control of the various functional disciplines involved in the project effort. Insufficient control and functional integration of project activities increases the risk of not achieving the project's cost, schedule, and performance objectives.

We note in Figure 3B that the Army and Air Force use the project team approach more frequently than the Navy, where it is used slightly more than 50% of the time. A plausible explanation is that, in general, when services are acquired at the installation level, the physical proximity of personnel can make it easier to establish and use the project teams in managing acquisition. Thus, the use of the regional approach in the Navy means that it has less opportunity to use the project teams. Perhaps a virtual team approach may need to be adopted.

Acquisition Leadership

In addition to the use of project teams, another best practice is formally designating a trained project manager/leader with the authority to lead the project effort. The project manager is typically a coordinator and integrator of the various functional disciplines involved in the project and has overall responsibility for the project's success. The project manager is focused on the overall objectives of the project, and integrating and balancing the interests of the various functional disciplines (engineering, procurement, finance, and logistics) involved in the service project. Figure 4 provides answers to the question: Who leads the services acquisition project, contracting officer (CO) or Quality Assurance Evaluator (QAE)/Contracting Officer Representative (COR)? Figure 4A shows that when a project team is used, the CO predominantly leads the services acquisition project in the Army and Air Force, and only slightly more than half of the time for the Navy. Figure 4B shows that when a project team is not used, the CO predominantly leads the services acquisition project in the Air Force and the Navy, and only slightly less than half of the time for the Army.

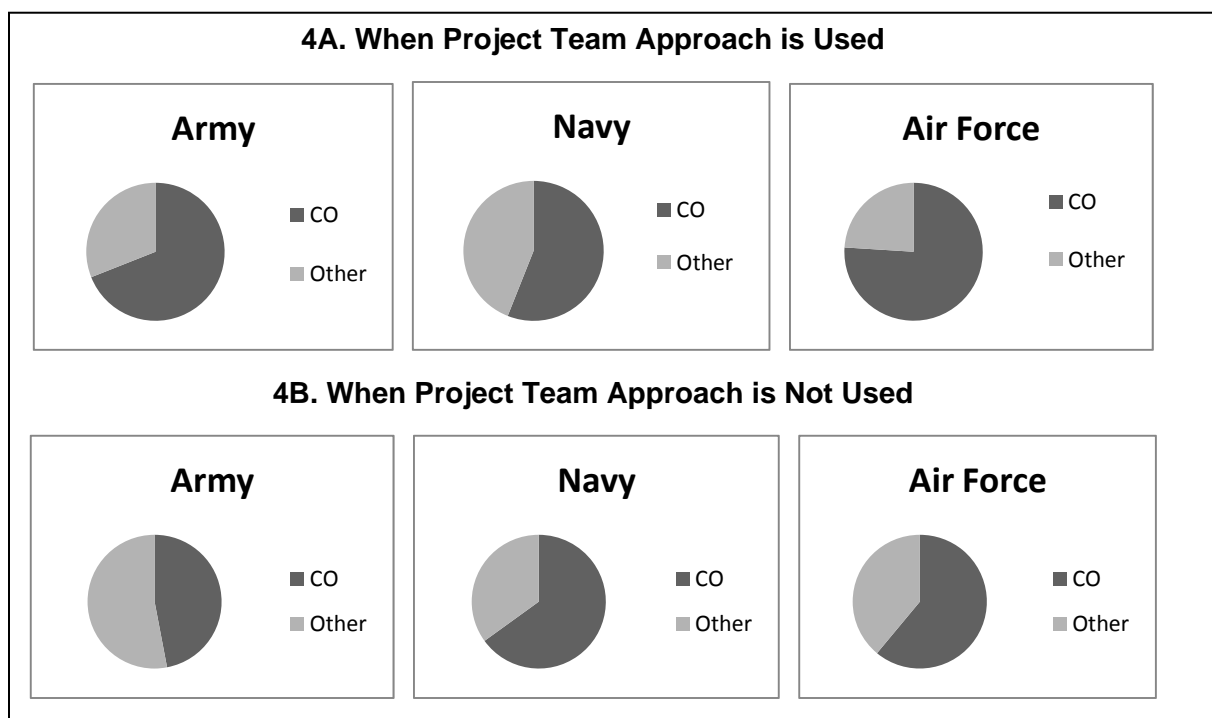


Figure 4. Acquisition Leadership

Requirement Ownership

Services acquisition includes managing the requirement. The “requirement” is the specific service that is being procured—for example, information technology services or aircraft maintenance services. It is important to note that the contract management process and, more specifically, the authorities and responsibilities of the contracting officer, do not include requirements management activities (such as determining the requirement, modifying the requirement, assessing the effectiveness of the requirement, or terminating the need for the requirement). These requirements management authorities and activities belong to the requirements manager of the organization responsible for the service being procured. Once the requirements organization identifies, develops, and defines the requirement, the contracting organization performs the contracting activities to procure the needed service. Contracting officers, however, may support the development of the requirements documents by providing business and procurement expertise in this area. For example, an aircraft maintenance squadron would own the aircraft maintenance service requirement being procured by the contracting organization for that specific installation. Figure 5 provides data on who owns the requirements, CO or QAE/COR.

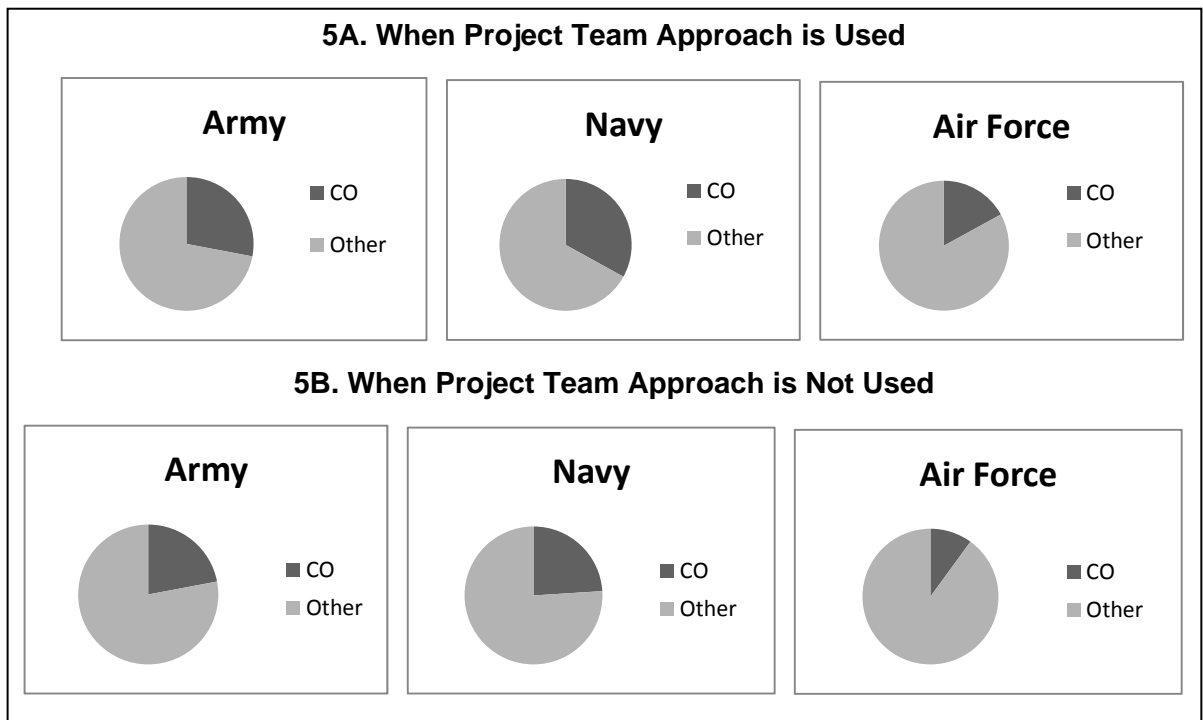


Figure 5. Requirements Ownership

In general, a contracting officer (CO) leading the acquisition, or owning the requirements, is not an appropriate practice, regardless of whether a project team approach is used or not. What is surprising from the survey data shown in Figure 4 is that the project teams are frequently led by the contracting officer as opposed to a formally designated project manager responsible for the overall service project success. As we note in Figure 5, less frequently, yet significantly, the contracting officer also owns the requirements. We consider these findings surprising since the contracting officer is a functional specialist concerned with ensuring the government rules and that the contract are in compliance by the contractor, while a project manager is concerned with the overall success of the project, in terms of cost, schedule, and performance objectives. In addition, a project manager typically represents the service requirement owner, and is typically authorized to make changes to the requirement during contract performance. Contracting officers do not have the authority to make changes to the service requirement, and, traditionally, do not have the expertise or technical knowledge to make such changes—for example, making changes to the requirements for aircraft maintenance service. Leading project teams involves managing the requirement and authorizing related technical changes to the requirement during contractor performance. We also observe the following in Figures 4 and 5:

- For the Army and Air Force, the use of a project team increases the probability of the CO leading the services acquisition;
- For the Navy, perhaps due to regional organization, the use of project teams decreases the probability of a CO leading the acquisition; and
- The above trends are the same for owing the requirements.

Program Management Issues

The first set of program management issues we investigated is the scope and ability of personnel responsible for services acquisition. Figure 6 provides comparative data on this count.

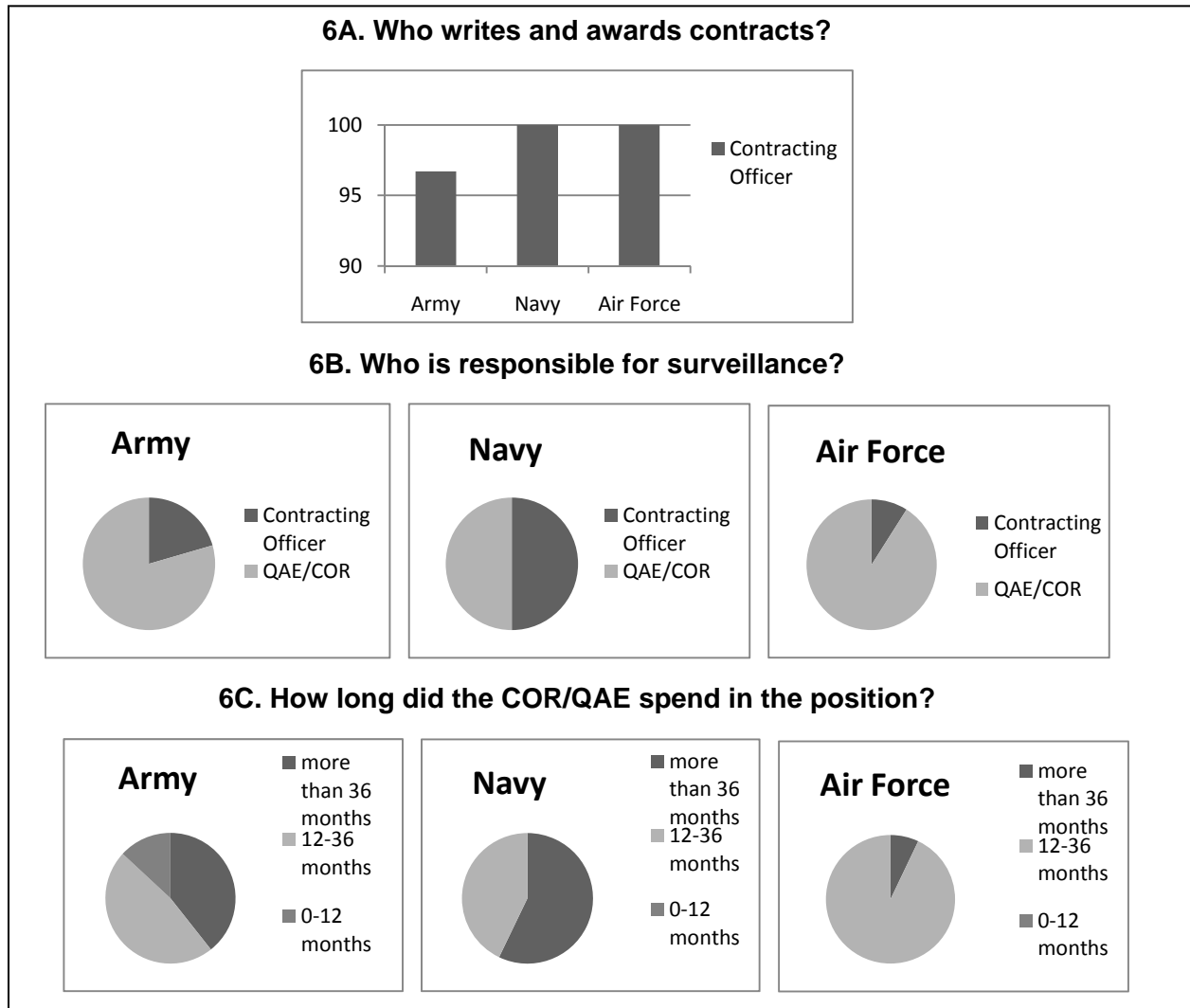


Figure 6. Scope and Ability of Personnel Responsible for Acquisition

Contractor Surveillance

As expected, we note in Figure 6A that the contracting officer always writes and awards contracts in the Navy and the Air Force. In the Army, the CO predominantly writes and awards the contracts. It is unclear why that would be the case. Another critical aspect of services acquisition is contractor surveillance. Contractor surveillance ensures that the contractor's performance complies with the requirements of the contract and, thus, the government is receiving the services procured. Due to the technical nature of government services contracts, contractor surveillance personnel should be knowledgeable of the technical aspects of the service, and are drawn from the technical community responsible

for the service requirement. Thus, having the requisite technical skills is critical for conducting contractor surveillance.

We note in Figure 6B that, as expected, in the Air Force and the Army, QAE/CORs predominantly provide contractor surveillance. However, in the Navy, QAE/CORs provide contractor surveillance in about 50% of the cases, with the contracting officer shouldering that responsibility in the remaining cases. These results indicate another situation in which contracting officers may be performing activities outside their area of expertise; in this case, performing contractor surveillance. Contractor surveillance involves technical knowledge and expertise in the service requirement area. A contracting officer, considered a business advisor with expertise in government contracting rules and regulations, should not be performing technical contractor surveillance on an aircraft maintenance service contract. Perhaps this is related to and caused by the regional approach to services acquisition being adopted by the Navy.

Finally, we studied the length of time COR/QAEs spend in their assigned position. The comparative data is presented in Figure 6C. We note that:

- The majority of COR/QAEs in the Air Force were assigned in the position for less than 3 years. Perhaps this is caused by significant turnover in staff.
- In the Navy, a significant percentage of COR/QAEs were in the job for more than 3 years. Interestingly, this seems to be the case in spite of the fact observed earlier that the contracting officer is responsible for surveillance half the time.

The final category of survey data consisted of other miscellaneous issues related to services acquisition program management. These include the use of lifecycle approach in routine and non-routine services, service acquisition billets, responsibility of various staff members, and the training they receive. The comparative data is presented in Figures 7 and 8.

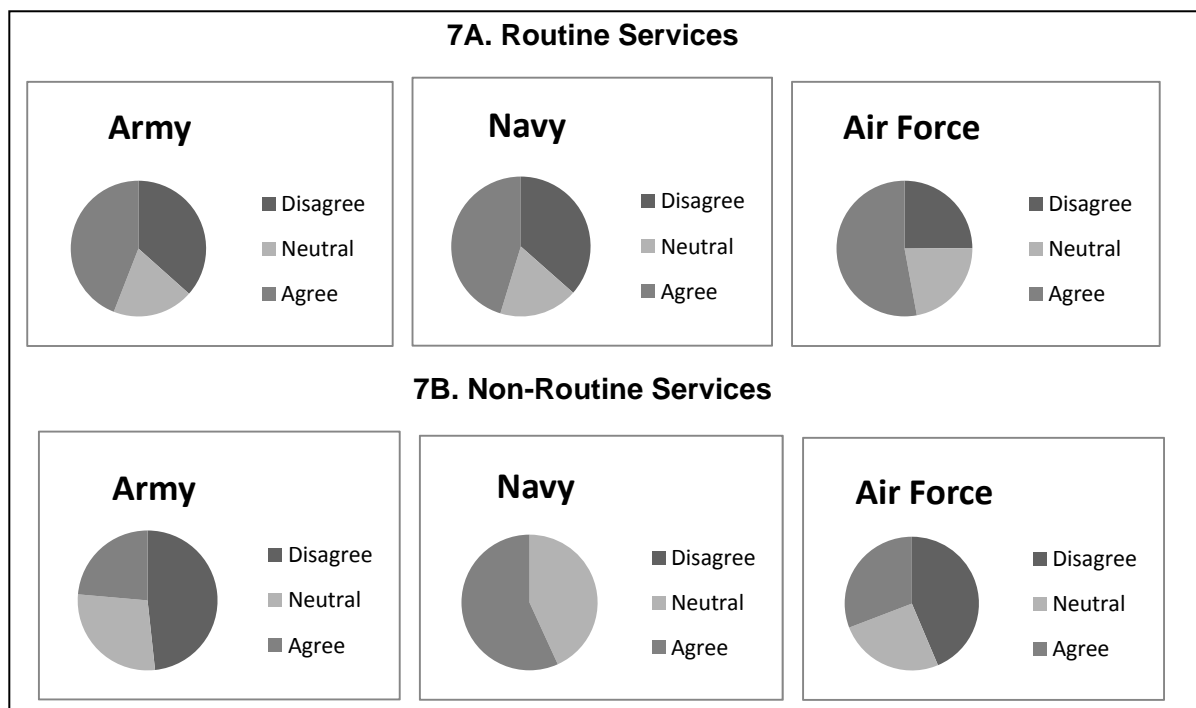


Figure 7. Lifecycle Approach

Lifecycle Approach

The use of a lifecycle to manage and control the progress of a project is considered a best practice in project management (Rendon & Snider, 2008). The project lifecycle allows the project to be managed in phases, with each phase controlled by gates and decision-points. The use of a project lifecycle should be a concern for ensuring proper management of service projects, especially non-routine services. If the services being procured and managed are of a non-routine nature, one would expect higher levels of uncertainty—and, thus, higher levels of cost, schedule, and performance risk—in the management of these service projects. Best practices in reducing project risk include the use of a project lifecycle—with project phases, gates, and decision-points for monitoring and controlling the progression of the service project procurement process as well as the resulting service. Without the use of a project lifecycle, the service project may be vulnerable to excessive risk in terms of meeting cost, schedule, and performance objectives. This would especially be true in the procurement and management of high-risk non-routine services.

Figure 7A reflects that, for routine services, a lifecycle was predominantly used by the Air Force, and less so (approximately less than half of the time) by the Army and Navy. As seen in Figure 7B, for non-routine services, a lifecycle approach was predominantly used by the Navy, and less so (approximately less than half of the time) by the Army and Air Force.

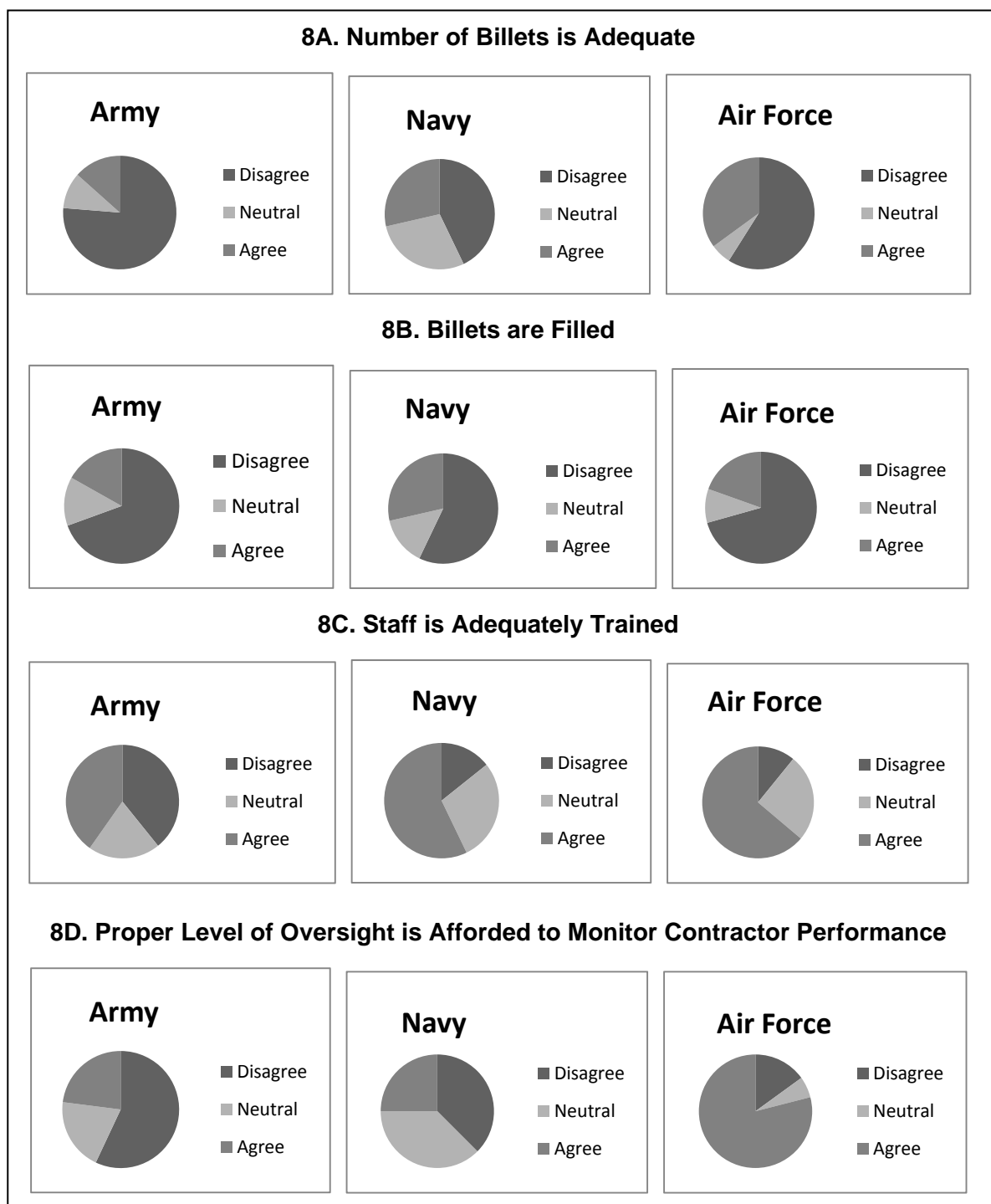


Figure 8. Acquisition Billets

Service Acquisition Billets & Responsibility of Staff Members

The management of services acquisition is the responsibility of the services acquisition personnel located at the regional or installation organizations. Each acquisition organization has designated acquisition positions, or billets, for the acquisition personnel. In

addition, these positions may or may not be filled, due to lack of personnel (perhaps personnel are deployed) or due to the understaffing of organizations. These acquisition personnel are also required to receive the appropriate training reflective of their assigned acquisition duties, such as contracting officer, quality assurance evaluator, or contracting officer representative. Thus, having an adequate number of acquisition billets in an organization is not sufficient. These billets must be adequately filled, and the personnel filling these acquisition billets must be adequately trained. Having an adequate number of filled acquisition billets, staffed with trained acquisition personnel, is integral for providing a proper level of oversight for monitoring contractor performance.

Proper level of oversight is integral to successful services acquisition management. Figures 8A through 8D reflect the survey responses for these areas. The following are salient observations on the figures:

- Figure 8A shows that the Army and Air Force predominantly disagree that there is an adequate number of acquisition billets, while the Navy survey responses are inconclusive.
- Figure 8B reveals that the Army, Navy, and Air Force all predominantly disagree that these acquisition billets are adequately filled.
- Figure 8C indicates that the Navy and Air Force predominantly agree that the services acquisition personnel are adequately trained, while the Army survey responses are inconclusive.
- Figure 8D suggests that the Army predominantly disagrees that a proper level of oversight is afforded to monitor the contractor's performance; the Air Force predominantly agrees that a proper level of oversight is afforded to monitor contractor performance, and the Navy survey responses are inconclusive.

Recommendations

To improve the management of services acquisition, our first recommendation is to maintain the positive trend of increasing the number of competitively bid, fixed-price contracts. These types of contracts promote competition, which ensures the government gets the right services at the best value. Fixed-price contracts shift the risk of cost overruns away from the government and onto the contractor. This also serves to incentivize the contractor to complete tasks within budget.

Our second recommendation relates to the management of services acquisition at the regional versus installation level. As previously discussed, each individual approach has advantages and disadvantages. In our view, the key to success under either approach is to use the proper supporting management processes. Consequently, we recommend that the Navy's regionalized management of services acquisition should implement a stronger project management approach, possibly a virtual project management team. This team would consist of the project manager, requirements manager, and contracting officer at the regional office. The QAE/COR would then serve as the site manager and be responsible for contractor surveillance. The QAE/COR would act as the "eyes and ears" of the regional project manager and contracting officer and would coordinate program and contracting issues back to the project manager. This may require QAE/CORs with higher-level knowledge and skills, due to their expanded roles and responsibilities. The Army and Air Force's installation-level management of services acquisition should ensure consistency in services acquisition management processes department-wide. Our recommendations include the establishment of dedicated installation project managers responsible for the



overall cost, schedule, and performance requirements of the services acquisition. Additionally, the installation project teams should include a requirement manager or representative that is authorized to identify, manage, and change the service requirement during the contract period. Establishing a dedicated project manager and adding a requirement manager/representative to the project team would relieve the contracting officer from performing these conflicting roles.

The third recommendation to improve the overall management of services acquisition is to increase the size of the acquisition workforce, reversing the downsizing trend that began in the 1990s. The results of this research show that the number of CORs/QAEs also needs to be increased. Increasing the size of the workforce will allow for better oversight and help ensure that contractor performance is properly monitored.

Our final recommendation is to increase the effectiveness and availability of training to ensure a qualified acquisition workforce. Based on the results from the research, a majority of respondents agreed that the acquisition workforce was adequately trained. Respondents also provided numerous negative comments regarding the poor quality and the lack of training. The recommended training should focus on all phases of the contract management process and related *Federal Acquisition Regulation (FAR)* policy. Additionally, training on areas related to working in cross-functional teams and using project lifecycles should be provided to all acquisition work force. Finally, and more importantly, if the contracting officers are to continue acting as de-facto project managers by leading the acquisition teams, then they should receive training on project management concepts, project control techniques, and project leadership.

Given the total amount of money spent and the scope of services acquisition in the Department of Defense, the opportunity for conducting research in this important area is limitless. One area that stands out as needing research is contracting for medical services. We have already started to address this need.

Finally, as discussed earlier, the researchers in the fields of operations management and marketing have studied and identified several key characteristics of services that lead to differences in the production and marketing of services as opposed to manufactured products. We believe that the same key characteristics also must be taken into account in designing and managing the processes involved in acquiring services. For example, *intangibility* of service outcomes makes it difficult to clearly describe and quantify services and, therefore, to contract for services. Intangibility of outputs also makes it difficult to define and measure quality. *Co-production*, requiring the presence and participation of customers in the creation of many services, is an important characteristic of services. Hence, the contracts for software development should ideally specify not only what the service provider should do but also what inputs the customer should provide. Otherwise, a satisfactory service outcome may not be realized. *Diversity* of services also makes it difficult and undesirable to use the same contract vehicles or procedures for different services. Finally, services are *complex* and may involve multi-stage processes. This makes it important yet challenging to write contracts that are flexible enough to cover all relevant scenarios and eventualities. Given these considerations, we believe that there exists significant opportunity to conduct research into the impact of these characteristics on the acquisition of various services and the associated implications for the management of service acquisition processes.



References

- Apte, A., Apte, U., & Rendon, R. (2008). *Managing the services supply chain in the Department of Defense: An empirical study of current management practices* (Technical Report NPS-AM-08-137). Monterey, CA: Acquisition Research Program, Naval Postgraduate School.
- Apte, A., Apte, U., & Rendon, R. (2009). *Managing the services supply chain in the Department of Defense: Empirical study of the current management practices in the Army* (Technical Report NPS-AM-09-136). Monterey, CA: Acquisition Research Program, Naval Postgraduate School.
- Apte, U., & Rendon, R. (2007). *Managing the service supply chain in the Department of Defense: Implications for the program management infrastructure* (Technical Report NPS-PM-07-126). Monterey, CA: Acquisition Research Program, Naval Postgraduate School.
- Apte, U., Ferrer, G., Lewis, I., & Rendon, R. (2006). *Managing the service supply chain in the Department of Defense: Opportunities and challenges* (Technical Report NPS-AM-06-032). Monterey, CA: Acquisition Research Program, Naval Postgraduate School.
- Camm, F., Blickstein, I., & Venzor, J. (2004). *Recent large service acquisitions in the Department of Defense*. Santa Monica, CA: RAND National Defense Research Institute.
- Compton, J.A., & Meinshausen, B.A. (2007). *The Department of Defense's management of services acquisition: An empirical analysis* (Master's Thesis). Monterey, CA: Naval Postgraduate School.
- Director, Defense Procurement and Acquisition Policy (DPAP). (2006, August 16). *Government Accountability Office high risk area of contract management*. Memorandum. Washington, DC: OUSD(AT&L).
- Director, Defense Procurement and Acquisition Policy (DPAP). (2007, March 2). *Contracts for services*. Memorandum. Washington, DC: OUSD(AT&L).
- DoD IG. (2009, April 22). *Summary of DoD Office of Inspector General audits of acquisition and contract administration* (DoD IG Report No. D-2009-071). Washington, DC: DoD.
- Ellram, L.M., Tate, W.L., & Billington, C. (2004). Understanding and managing the service supply chain. *Journal of Supply Chains Management*, 40(4), 17-36.
- Federal Procurement Data System (FPDS). (2010). Retrieved January 7, 2010, from <https://www.fpds.gov/>
- Fitzsimmons, J.A., & Fitzsimmons, M.J. (2006). *Service management: Operations, strategy, and information technology* (5th ed.). New York: McGraw-Hill.
- GAO. (2001, May). *Contract management: Trends and challenges in acquiring services* (GAO-01-753T). Washington, DC: Author.
- GAO. (2002a, January). *Best practices: Taking a strategic approach could improve DOD's acquisition of service* (GAO-02-230). Washington, DC: Author.
- GAO. (2002b, July). *Acquisition workforce: Agencies need to better define and track the training of their employees* (GAO-02-737). Washington, DC: Author.
- GAO. (2005, March). *Contract management: Opportunities to improve surveillance on Department of Defense service contracts* (GAO-05-274). Washington, DC: Author.
- GAO. (2007a, January). *Defense acquisitions: DOD needs to exert management and oversight to better control acquisition of services* (GAO-07-359T). Washington, DC: Author.



- GAO. (2007b, May). *Defense acquisitions: Improved management and oversight needed to better control DoD's acquisition of services* (GAO-07-832T). Washington, DC: Author.
- GAO. (2009a, January). *High-risk series: An update* (GAO-09-271). Washington, DC: Author.
- GAO. (2009b, March). *Department of Defense: Additional actions and data are needed to effectively manage and oversee DOD's acquisition workforce* (GAO-09-342). Washington, DC: Author.
- GAO. (2009c, April). *Defense acquisitions: Actions needed to ensure value for service contracts* (GAO-09-643T). Washington, DC: Author.
- Gordon, G.L., Calantone, R.D., & Di Benedetto, A. (1993). Business-to-business service marketing: How does it differ from business-to-business product marketing? *Journal of Business and Industrial Marketing*, 8(1), 45-57.
- Hutt, M.D., & Speh, T.W. (1998). *Business marketing* (6th ed.). New York: Dryden Press.
- Lovelock, C.H. (1992). Are services really different? In C.H. Lovelock (Ed.), *Managing services* (2nd ed.) (pp. 1-8). Englewood Cliffs, NJ: Prentice-Hall.
- Metters, R., King-Metters, K., & Pullman, M. (2003). *Successful service operations management*. Mason, OH: South-Western.
- Rau, C.A., & Stammersky, P.J. (2009). *Management and oversight of services acquisition within the United States Army* (Master's Thesis). Monterey, CA: Naval Postgraduate School.
- Rendon, R.G., & Snider, K.F. (Eds.). (2008). *Management of defense acquisition projects*. Reston, VA: American Institute of Aeronautics and Astronautics.
- Schiele, J.J., & McCue, C.P. (2006). Professional service acquisition in public sector procurement. *International Journal of Operations and Production Management*, 26(3), 300-325.
- Smeltzer, L.R., & Ogden, J.A. (2002). Purchasing professionals' perceived differences between purchasing materials and purchasing services. *Journal of Supply Chains Management*, 38(1), 54-70.
- US Department of Labor. (2005). *Archived News Releases for 2005; Employment Situation*. Bureau of Labor Statistics. Retrieved from <ftp://ftp.bls.gov/pub/news.release/History/empsit.02042005.news>

General Services Administration Streamlines the Procurement of Construction Services

Jeffory Meyer—Meyer is a project manager with the US General Services Administration, Region 6 and is also a Team Leader for the West Project Management Branch. Jeff is a licensed Architect, NCARB Certified and a member of CIB. Meyer has been recognized for his numerous accomplishments for incorporating environmentally conscious initiatives into his projects. He is the author of a research paper regarding highly insulated roofs and a co-author for a case study of non-prospectus projects for the GSA. Prior to GSA, Meyer was a practicing Architect for 15 years and a Principal of an Architectural firm for 6 years.

Jeffory E. Meyer, NCARB, RA
Project Manager
Design & Construction Division
US General Services Administration, Region 6
Public Building Services
Kansas City, MO 64131-3088 USA
Phone: 816-823-2260
Mobile: 816-564-2339
jeff.meyer@gsa.gov
International Council for Research and Innovations in Building and Construction (CIB)

Stephanie Witt—Stephanie Witt is a contracting officer with the US General Services Administration, Region 6, Design and Construction Division and is also a Team Leader for the West Acquisition Branch. Witt has been recognized for her contracting efforts and knowledge and is looked to as a highly knowledgeable source regarding the FAR and Contract Requirements within the GSA. She has served on many committees dedicated to determine the proper way to implement contracting requirements. She is known for her willingness to test and modify various processes that have not been previously used within the GSA.

Stephanie R. Witt
Contracting Officer
Design & Construction Division
US General Services Administration, Region 6
Public Building Services
Kansas City, MO 64131-3088 USA
Phone: 816-823-5013
Mobile: 816-215-4142
stephanie.witt@gsa.gov

Jacob Kashiwagi—Jacob Kashiwagi is a program manager at PBSRG and a lecturer at ASU. He is the developer of the no-influence leadership theory, the theoretical basis for the PIPS model. The technology has been tested over 600 times totaling \$2.4 billion (\$731 million in construction and \$1.7 in non-construction projects) with a 98% success rate since 1994. Kashiwagi is also the lead researcher for the information model implemented at US Army Medical Command (\$400 million in construction renovation awards per year) that forced the contractors to concentrate on value and not price. He is also an author of several research papers and reports.

Jacob Kashiwagi, MS, Researcher
Arizona State University, School of Sustainable Engineering and the Built Environment, Performance Based Studies Research Group (PBSRG)
Tempe, Arizona USA
Phone: 480-965-4273
Email: jacobk@exchange.asu.edu



Dean Kashiwagi—Dean Kashiwagi is a professor at Arizona State University's Del E Webb School of Construction and also the Director of the PBSRG. He is the recipient of the distinguished 2009 Excellence Educator Award for his numerous accomplishments in education and research. His many achievements include receiving a Fulbright Scholar award to share state-of-the-art facility and project management research and practices with the people of Botswana, Africa. His groundbreaking best value PIPS and PIRMS Model was integrated into a graduate program. Prior to ASU, Kashiwagi was a Project Engineer for the US Air Force during his 14 years of military service.

Dean T. Kashiwagi, PhD, PE, Director
Performance Based Studies Research Group (PBSRG)
Arizona State University
Tempe, AZ 85287-0204
Phone: 480-965-4273
Email: Dean.kashiwagi@asu.edu

Abstract

The General Services Administration (GSA) Heartland Region is implementing a best value process (which minimizes time and cost deviations 98% of the time) and which minimizes the need for client's decision making, transfers the risk and control of a project to the vendor, and forces the vendor to manage and minimize the non-technical risk that the vendor does not control. The Performance Information Risk Management System (PIRMS) has been tested by the US Army Medical Command and has minimized over 50% of client project management and risk management transactions, and also minimized cost and time deviations by as much as 70%. The new paradigm uses Deming's concept of managing and minimizing the project deviation instead of meeting minimum standards. The system forces the client's representatives to do quality assurance, and the vendor to do quality control. The mechanism used is a risk management plan and a weekly risk report that creates transparency between buyer and vendor. The system can minimize up to 90% of the government's transactions and activities. The system is a new paradigm for government systems.

Keywords: Best value procurement, minimized government management, high vendor performance, and measured environment

Introduction

The General Services Administration (GSA) is the largest buyer of non-military services in the United States. It is a large management based organization. An Achilles heel for any large organization is the number of layers of management, the large number of managers and subject matter experts (SME) and the practice of managing, directing and controlling vendors/contractors who are supposed to be experts at what they do.

The current status of most projects in the GSA is where vendors (architect/engineers and contractors) continually rely on being managed, directed, and controlled by GSA project managers and contracting officers. To get a quality set of construction documents, the government project managers complete extensive quality control reviews of the A/E's construction documents, once the sole responsibility of the A/E design firms. GSA personnel (CORs, PMs, and COs) continually manage, direct, and control the contractors in construction. Control and risk are not transferred to the vendors, thus making it difficult to hold vendors accountable for deviations. Projects incur change orders due to design deficiencies. Projects are not being completed in a timely manner and the actual close-out of projects could take between 1 to 4 years. GSA processes and requirements are



continually being developed and expanded at various levels both nationally and regionally in an attempt to increase the performance of the vendors.

Upper level management in the GSA has struggled with implementing a sustainable, useable, and accurate system that measures the performance of their vendors and project managers. Shrinking budgets, increased workload requirements, and the increased need for project managers to manage, direct, and control vendors, make the updating, collection, and analysis of performance information very difficult. The GSA has been exposed to many management measurement systems and philosophies (Alsup, 2010; Topi, 2010):

1. Quality Management Circle (part of TQM)
2. TQM (total quality management) (early 1990s)
3. eTMP (electronic transaction management playbook) (2006)
4. TMP (transaction management playbook) (2006)
5. HCAM (included the following TMP, OMP, LCP, & AMP) (2005)
6. OMP (occupancy management playbook)
7. LCP (large construction playbook)
8. AMP (account management playbook)

However, the strategic objective of increasing performance and value of vendor services and measurement of the performance remains elusive.

Problem

The stubbornness of the problem of the GSA's inability to sustain performance measurements in a timely fashion and increase vendor performance may be a systems problem and not a GSA unique management/leadership problem. The current GSA system forces the project managers to document, maintain, and report the performance information. Because of the current project manager/vendor relationships and their heavy workload, project managers may not be motivated to accurately and consistently document the performance information. The current system of delivery has the following suboptimal characteristics:

1. The GSA project managers are required to manage, direct, and control the vendors.
2. There is no transfer of risk or control to the vendors.
3. The relationship between the vendors and the project managers may dilute accountability.
4. The current delivery system does not motivate vendors to preplan and manage and minimize the risk that they do not control (think in the best interest of the client).

Hypothesis

Deming (1982) identifies the GSA problem as a systems problem. The authors propose that the system may be stabilized but not meeting the expectations of the GSA's upper level management. Increasing effort to optimize performance in a stable environment may not be successful. The system must be changed to increase performance, and have a



sustainable performance measurement system that results in an increase in vendor performance.

Methodology

The authors propose that the GSA find a new system that has the following characteristics:

1. Run by a core team of systems managers who understand a performance information based system.
2. Selects vendors by their capability to understand a project and manage and minimize deviations that are caused by sources outside of their control.
3. Transfer the risk and control to vendors, who by contract must preplan, manage and minimize risk that they do not control, and manage the risk of their projects by measuring and minimizing cost and time deviations.
4. Measure the vendors' performance.
5. Vendors are held accountable for all deviations unless they can dominantly document the source of the risk and how they attempted to manage and minimize the risks.
6. System has past performance of success in other government organizations.

Search for Performance Measurement Systems for Vendors/Organizations

The GSA team composed of a project manager and a procurement officer proposed the above plan to their division manager in the heartland region located in Kansas City. They proposed using a best value approach to solicitation to procure the services of an organization that could:

1. Provide a system that meets the requirements of the new system.
2. Show documented evidence that the system measured performance and resulted in a dominant increase of performance.
3. Show capability to educate and train GSA personnel to be able to understand the new system and run the system.

The solicitation was posted on August 15, 2009. There were only three proposers. The GSA identified the Performance Information Procurement System (PIPS) as the only system that could meet the new requirements. It was the only system that:

1. Had documented performance over 15 years.
2. Minimized the amount of management, direction, and control.
3. Transferred risk and accountability to the vendor.
4. Measured the vendor and all other participants who interfaced with the vendor.
5. Provided a mechanism whereby the vendor managed the risk that they did not control.
6. Minimized the transactions (meetings, emails, and telecom.)



No other vendor options would accept responsibility for ensuring the performance measurements on all participants, a resulting increase in performance and value, and minimized management, direction and control. The other vendor options proposed to put a performance measurement system in place, but would take no responsibility for the successful implementation or the impact on performance of the organization and the vendors. The best value PIPS system dominantly differentiated itself as the only option with the potential of disruptive change required to employ a sustainable measurement system and result in a dominant increase in performance and value.

Industry Analysis

The construction industry structure (CIS) model (Figure 1) (Kashiwagi, 2009) identifies the difference between what the GSA and other owners currently employ, and what is needed to employ a sustainable performance measurement system that has accompanying increase in performance and value.

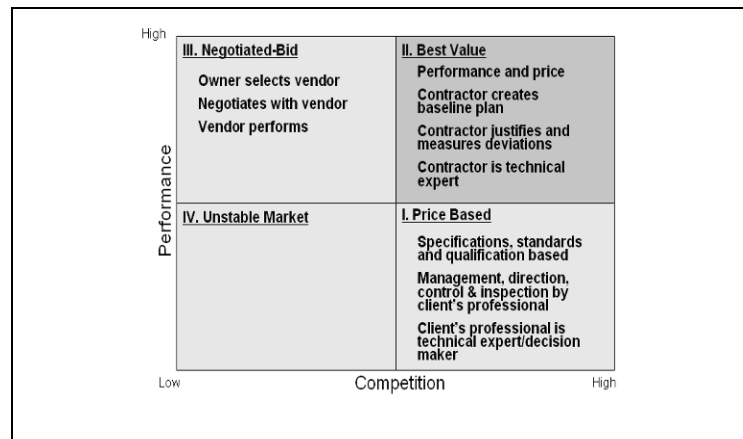


Figure 1. Construction Industry Structure
(Kashiwagi, 2010)

The CIS identified the current GSA system as a price based system regardless of the perception of not awarding projects based on price (Sullivan, 2005). The Price Based System (Quadrant I), has the following characteristic:

1. Owner representative attempts to direct and control the vendor.
2. Deductive logic assumes that the owner would only hire a vendor who knew what they were doing (expert).
3. Someone who should know less about what is being done is directing someone who is an expert in what is being delivered.

In the best value system, the client identifies their intent, but asks the industry (vendors) who can deliver the best value that meets the intent of the owner? This system has the following characteristics:

1. Vendors identify what they do, and what it costs.
2. Client selects the best value based on performance and price (value).
3. Vendors preplan, manage and minimize risk they don't control. (They have no technical risk; they are experts.)

4. Clients do quality assurance (ensure that the vendor has quality control and risk management systems).

The price based system is:

1. Inefficient.
2. Requires more people.
3. Has more confusion.
4. Subjective.
5. Higher flow of information required between parties.
6. If performance measures are kept, they are very subjective and potentially could be construed as biased. There are more measurements that require an expert to decide what is good performance.
7. Requires people to partner.
8. Relationships are important between vendors and the client's PMs.
9. Adversarial.
10. Because minimum requirements are used, performance will be in decline.
11. Not accountable.
12. Has high risk.
13. Low performance (on time, on budget, meet client's expectations.)

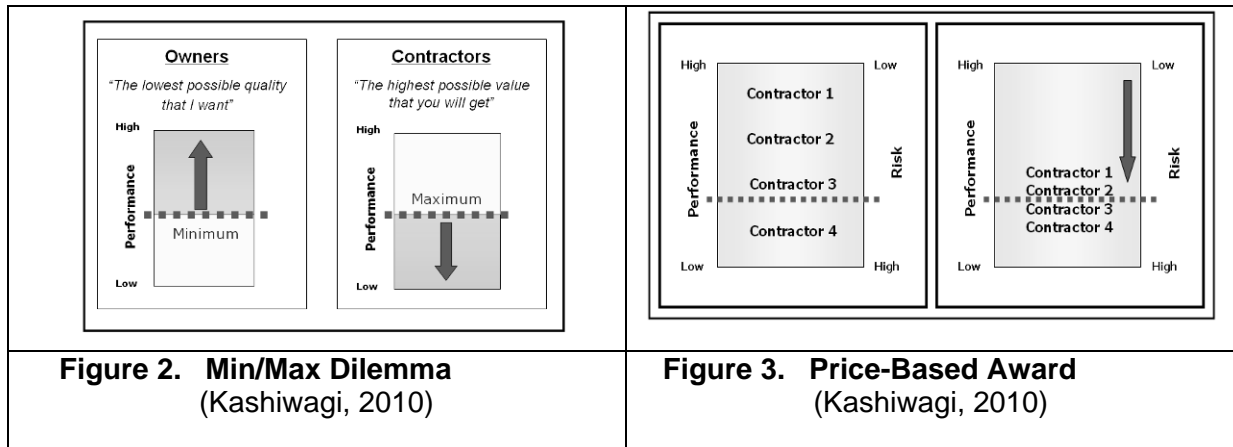
The best value system:

1. Is more efficient.
2. Requires less people.
3. Is measured with few and simple measurements.
4. Is transparent.
5. Uses alignment of resources.
6. Has minimized flow of information.
7. Minimizes partnering exercises.
8. Managed by minimizing deviations.
9. Has continuous improvement.
10. Has a high level of accountability.
11. Has high performance and low risk.
12. Minimizes transactions.



Difference Between the Two Systems

The difference between systems is that the price based vendor is reactive, managed, controlled and directed by the client/buyer; the best value vendor is proactive, preplans, manages and minimizes risk that they do not control, measures their performance, and manages their project by minimizing deviations. In the price-based sector, the client directs the vendor using minimum requirements; the vendors transform the minimum into a maximum, and drive the performance the opposite direction (Figure 2).



The high performance vendor (high performance and low risk) can see a project from beginning to end. They identify risk that they do not control, and plan solutions to minimize the risk. The low performer (high risk) prices only what they are directed to price. The overall effect of the client directing the vendors is the following (Figure 3):

1. High performers become reactive instead of proactive. They are told to price only what is directed, regardless of completeness, correctness, or whether it is doable.
2. They are directed to assume that the directions are perfect.
3. Therefore, they are directed to give the lowest possible price, assuming nothing goes wrong.

Figure 4 shows the business approach to Quadrant I Price Based system. The high performers, who get paid more for their expertise, get sent to the price based system, where they are directed and controlled by someone who is not an expert. The confusing environment results in lower production of the high performer. They become out of alignment, overpriced, and leave the environment. A more damaging result of the system is that the less experienced are not motivated to become like the highly trained.

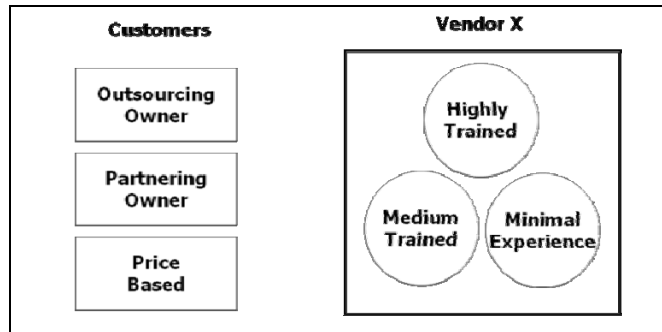


Figure 4. Business Approach
(Kashiwagi, 2010)

The outsourcing owner who transfers risk and control to the expert vendor will get the high performer. It is the only win-win situation. The high performer will preplan, manage and minimize risk that they do not control to finish on time and maximize their profit, do the project once, and get paid and not go back to redo or fix problems. This is the efficient best value system.

Best Value Performance Information Procurement System (PIPS)

After identifying the requirement to have a new system/environment that reflected the best value environment, the GSA selected the best value PIPS system to move from the price based to the best value environment (GSA contract #: GS06P09GYD0027). The documented performance of the PIPS system included:

1. 16 years of testing (1994-present) delivering 700+ construction services projects valued at over \$800 million.
2. Research funding of \$8.5 million.
3. Minimized client risk/project management activities by up to 90%.
4. Maximized vendor profit by up to 100%, at no additional cost to the client.
5. Delivered performance of 98% on time, no contractor generated cost and time deviations, and meeting client's expectations.
6. Arizona State University (ASU) moved PIPS into non-construction areas including the delivery of food services, IT networking, IT data centers, help desks, sports marketing, gym equipment, document control, long distance education services, and furniture buys. ASU received investments of \$100 million over ten years due to the change of system environment, from price based to best value.
7. The Dutch infrastructure agency used PIPS to deliver \$1 billion of highway infrastructure to solve their problems with the delivery of construction.
8. The Bank of Botswana used PIPS to deliver a critical bank facility and found it tremendously better than the traditional process.
9. The State of Alaska is delivering a \$200 million Electronic Resource Planning system using PIPS.

PIPS is a licensed structure/process from Arizona State University developed by Dean Kashiwagi and the Performance Based Studies Research Group (PBSRG.) PIPS has three main phases (Kashiwagi, 2010):

1. Phase I: Selection of the best value vendor.
2. Phase II: Pre-award/pre-planning and creation of the risk management plan (RMP) and the weekly risk report (WRR.)
3. Phase III: Project delivery by the risk management of deviation of time and cost.

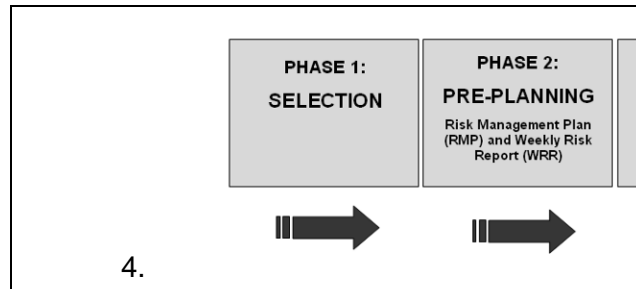


Figure 5. PIPS/PIRMS Phases
(Kashiwagi, 2010)

During the early development of PIPS, the selection phase was identified as the most important phase. As more tests were run, it was identified that the pre-award, preplanning phase was a more critical phase. The risk management capability of PIPS became obvious, and the term, the Performance Information Risk Management System (PIRMS) was created to allow owners to use the risk management capability of the system even though they would select their vendor using the price based selection mode.

PIPS has six major filters in the selection phase to ensure performance (Figures 6 and 7):

1. Requires the vendors to prove the potential to perform through documented past performance including the performance of critical personnel (project manager), and critical sub-vendors (engineering, professional consultants, or other crafts).
2. Capability to do the project. This submittal is a blind review of the vendor's capability to do the project:
 - a. Requires the vendors to show an understanding of the scope of the project in terms of the biggest technical risks in a concise and short explanation (two page submittal). The high performers should identify the major technical challenges of the project and what makes them capable of minimizing the risk of nonperformance.
 - b. Risk Assessment/Value Added (RAVA) submittal that forces the vendor to identify the risk that the vendor does not control, and how they will manage and minimize that risk so it does not occur. It also asks the vendors to document dominant added value being offered by the vendor (that is not in the client's specified scope, and will create a dominant

difference in project value) that makes them different from their competitors.

- c. The vendor is also asked for a milestone schedule, and how they will measure their performance during the project.
 - d. The vendor will also submit a cost breakout.
- 3. The interview of the critical personnel of the vendor to identify the person's relative vision, ability to predict things before they occur, preplan and their capability to be accountable.
 - 4. A cost check to ensure the best value is not overly expensive.
 - 5. Prioritization of the best value based on the capability to perform (the past performance information, scope rating, RAVA rating, milestone schedule rating, rating on performance measurement system, interview rating and price).
 - 6. Pre-award Phase where just the best value vendor creates a risk management plan (RMP) and a weekly risk report (WRR) that they will use to manage and minimize the deviation of the project. The WRR and RMP track risks that the vendor does not control. This is the key mechanism in PIPS/PIRMS, and is the regulator that ensures that risk and control is transferred to the vendor. This becomes a key component of the contract, and decommissions any attempts for the owner's PM to manage, direct, and control the vendor. This results in an alignment of resources, as it is in the best interest of all parties.

A major paradigm shift is the movement from management, direction, and control to quality control/quality assurance. This movement assumes that the vendor has no technical risk, and therefore will concentrate on managing and minimizing the risk they don't control. This will be more fully explained in a following section.

The vendor then writes their own contract (technical requirement, legal requirements of the owner, risk management plan and weekly risk control report.) The project is then awarded. The vendor self manages themselves based on the contract, managing and minimizing the cost and time deviation of the project. At the end of the project, the vendor is rated. The rating becomes 50% of the vendor's future rating.



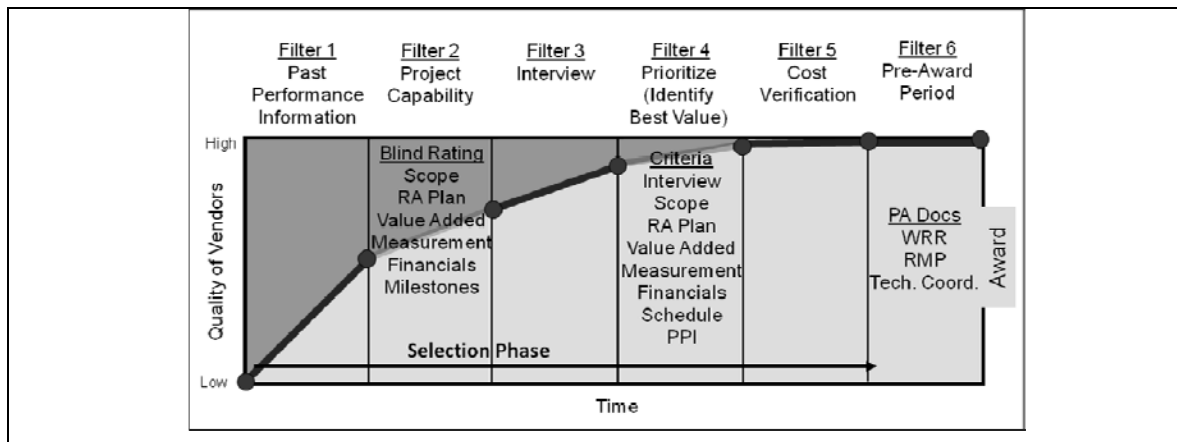


Figure 6. PIPS/PIRMS Filters
(Kashiwagi, 2010)

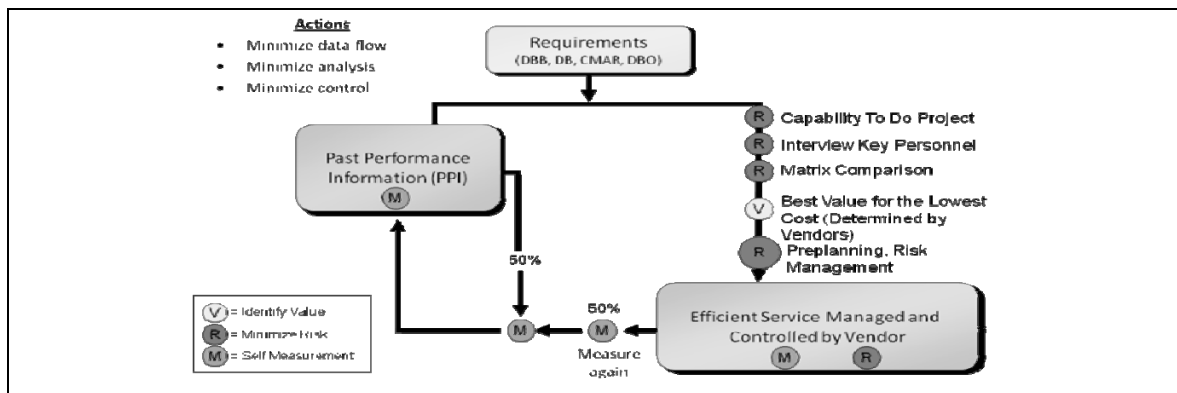


Figure 7. PIPS/PIRMS Self Regulating Closed Loop
(Kashiwagi, 2010)

A Paradigm Shift: Understanding that Expert Vendors Have no Technical Risk

A major departure from the traditional project management practices is the understanding and handling of risk. Information Measurement Theory (IMT) identifies that by definition, high performance/expert personnel have minimal or no technical risk (Kashiwagi, 2010). If there is technical risk, it is only because the client hired a vendor who does not have the expertise and therefore is not capable of minimizing risk. Instead of managing, directing, and controlling the vendor, the owner is now creating a new environment, where the vendor is identified as the expert. Therefore, the new environment minimizes all management, direction, and control of the vendor. This simple but difficult change in paradigm, forces the transfer of risk and control to the expert, and aligns all resources.

The authors propose that the impact of unforeseen conditions can be minimized if experts are used to manage and minimize the risk that they don't control. The only risk high performers have is risk that they do not control (risk that is brought by other participants,

mainly the client in the form of over-expectations, items outside of the scope, decision making by other participants at the wrong time during the process, and the changing of expectations) (Figure 8).

The new paradigm motivates vendors to preplan the project from beginning to end and know the risk that they do not control and how they will manage that risk before they accept the project. By deductive logic, a system that increases client management, direction, and control moves the activity to the more inexperienced vendors and personnel (Figure 8). This results in lower performance, reactive behavior, minimum standards or expectations, and minimum accountability.

Price based contracts emphasize the technical risk that the vendors must control. Price based contracts attract the less experienced, and makes the very experienced less competitive (Figures 4 and 8). Best value contracts must identify and communicate the expectations of the client but emphasize the requirement of the vendors to manage and minimize the risk that they do not control, thus thinking in the best interest of the client and creating a “win-win.” Price based contracts must cater to the inexperienced and increase the flow of information, contract documents, and client management, direction, and control. Best value contracts cater to the high performing contractors who need minimal information, who act in the best interest of the client by giving high technical service (no technical risk) and manage and minimize the risk that the vendor does not control through the use of quality control and risk management plans.

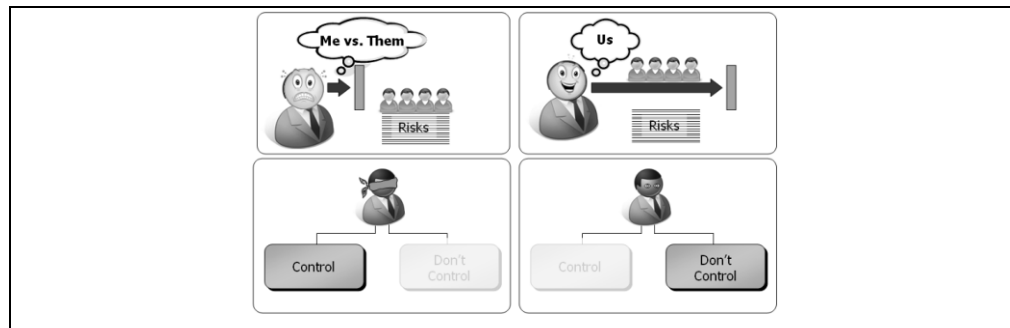


Figure 8. Inexperienced versus Experienced Vendor Risk Model
(Kashiwagi, 2010)

High performers and experts by definition are rarely surprised, are rarely affected by “unforeseen conditions,” and if there are “unforeseen conditions,” high performers have a preplanned solution that they can easily identify, manage, and minimize the cost and time deviation. Due to the low performing mentality of the construction industry, and resulting low performance of the industry participants, the risk of “unforeseen conditions” has increased, and instead of requiring high performance and the minimization of risk that the vendor does not control, the industry has created a transaction based solution of client management, direction and control, which increases the number of participants who do not have the capability to minimize the risk. These participants create silos where they verify and manage the vendors schedule and cost, negotiate prices, and approve all deviations in schedule or materials. In this inefficient system, the inexperienced contractor becomes more competitive than the high-performance contractor with very experienced personnel, who cannot efficiently do their project in an environment of redundant and un-needed transactions.

These deductive concepts have not only been confirmed through the 15 years of testing, but have been identified in other non-construction industries. By minimizing the

management, direction, and control of subject matter experts (SME), transferring the risk and control to the vendor, by forcing the vendor to write the contract which requires an expert vendor, and by forcing the vendor to manage and minimize the risk they do not control, a structure has been created which aligns resources and minimizes the transactions. The regulator of the structural change is the weekly risk report (WRR) and the risk management plan (RMP.) They create transparency, which minimizes the need for management, direction and control.

Method of Measurement

The method of measurement of performance utilized by PIPS is a novel approach which has had successful results in the past five years (Kashiwagi, 2009). The assumption is that a large organization delivering services with project managers who are accustomed to manage, direct, and control, will have a difficult time consistently and doing timely reporting and analysis of performance due to the inefficiency of a management based system. PIPS identifies who should be at risk (expert), and forces the measurement of performance (deviation of project time and cost) by the vendors. The deviations are then reviewed for accuracy by the client/buyer's professional (quality assurance.) The information goes directly to the top decision maker in the organization, bypassing the normal filtering system in a bureaucracy. A simplistic use of spreadsheets does the following:

1. Identifies the deviation rates of all projects.
2. Identifies the top ten riskiest projects, and which participant in the supply chain caused the deviation.
3. Measures all participants in the supply chain, giving a relative performance rating based on deviations and performance.

The PIPS measuring system overcomes the major obstacles large organizations have:

1. Subjective filtering of measurements.
2. Lack of timely reporting.
3. Lack of time to do accurate reporting.

GSA Strategic and Tactical Plans to Implement New Environment

The GSA signed a five-year contract with the Performance Based Studies Research Group (PBSRG) to implement the new best value system. Previous research results (Kashiwagi, 2010) identified that both a strategic plan and a tactical plan are required to successfully make the transformation. Previous results identified the following priorities:

1. Identification of a core group of visionaries who could understand the theoretical change in systems, and be systems managers and managers of personnel.
2. Education, development of the core group of visionaries to learn how to use the system.
3. Implementing and modifying PIPS to fit the environment of the owner.
4. General education to the owner's organization of the PIPS system.



The strategic plan encompasses the first two objectives, and the tactical plan is the second two objectives. The tactical plan (second two objectives) cannot be implemented without the strategic plan. The strategic plan therefore, must include the following:

1. The development of the visionaries.
2. Theoretical education of the visionaries.
3. Job transformation of the visionaries from project managers to educators and systems managers.
4. Changing of the effectiveness of the chain of command.
5. The visionary group must learn to apply the concept of PIPS to their organization's transformation.
6. Documentation of the transformation by the system.
7. Peer review by other visionaries located at other organizations who are working toward the same goals.
8. Development of measurements and a schedule showing the improvement in performance and value.

The tactical plan must include:

1. The modification of the PIPS to move the owner's organization without increasing resistance due to the change of efficiency and structure.
2. Education on PIPS to both vendors and PMs running the system.
3. Prototype testing and implementing PIPS by core team visionaries.
4. Design of the information system.

If the concept of transferring risk and control to the expert vendor, and aligning the resources in the entire supply chain through measurement of deviation is accurate, this system transformation is not industry specific. It is a system regulated by measurement, which aligns all the participants in the supply chain, minimizing the transactions and forcing experts to be accountable. It will bring discomfort if a subject matter expert (SME) is misaligned or does not currently have accountability. Therefore the transformation has to be "gentle" and evolutionary.

Progress of the GSA

The implementation of PIPS in the GSA has been relatively optimized due to the following:

1. The head of the core team visionaries, the region director of the organization, was already attempting to transform the organization to a measured organization. His strategic goals of efficiency and effectiveness of both vendors and the GSA organization was already in place.
2. The PMs of the core team were identified and selected based on the Information Measurement Theory (IMT) and therefore were attempting to use the concepts of PIPS before the transformation effort.



3. This is the first time in 16 years of testing that both a PM and procurement officer were original members of the visionary core team, and the director was already attempting the transformation.
4. This is the first time that the PIPS was selected through application of PIPS, thus confirming to the core team that PIPS was dominant in its ability to transform organizational environment/systems.

First Measurement of Existing Performance

The core team selected 8 projects where information was readily available to identify the performance of the existing environment. The performance measurements include:

1. Average cost/scope of projects: \$526,992
2. Average duration of projects: 152.5 days
3. Cost deviation of projects (percentage): 7%
4. Time deviation of projects (percentage): 231%
5. Customer satisfaction (1 -10 rating, 10 being optimal): 6.5

The core team is also interested in the following measurements that will be collected through surveys:

1. Vendor profit margin: TBD
2. Vendor rating of delivery system: TBD
3. Vendor perception of new system (1 – 10 rating): TBD
4. Number of projects a PM is responsible for: TBD

The GSA's next step is to complete six (6) test projects and collect data to confirm increased performance with the new best value PIPS system. The Contractor and PM shall rate the following before and after on the traditional system vs. the new best value PIPS system:

1. Effectiveness (deviations.)
2. Value of preplanning by vendor as perceived by both the vendors and the GSA PM.
3. Value of vendor managing and minimizing the risk that the vendor does not control as perceived by all participants.
4. Vendor's profit margin maximization.
5. Accountability of all the participants as perceived by all participants.
6. Successfulness and impact of the transfer of risk and accountability to the vendor.
7. Project coordination by the vendor with the client.
8. Minimization of surprises.

Schedule of Implementation

It is a five-year tactical plan:



1. Year 1: set up core team structure. Run the first tests with core team and a few PMs.
2. Year 2: set up the Directors Report and expand both the running of the entire process and the risk management reporting (which measures the performance of projects).
3. Years 3 - 5: expand implementation within organization. Visionary core team becomes a subject matter expert (SME) to assist in the transformation of other organizations.

Analysis of the Effort

This research effort is using the deductive approach (confirmatory) instead of the inductive approach (exploratory.) The success of the project will be determined by measurements of observation which minimize subjectivity as much as possible. The following are observations of the effort thus far:

1. PIPS has been identified by a GSA selection process as the only option with documentation of proven success to transform an organization's environment from a management, direction, and control environment to a best value, alignment, leadership based environment.
2. A large federal organization who is constrained by federal law, will implement the PIPS process for selection of vendor, and contract administration.
3. A visionary core team has been organized that is optimal in terms of a high-ranking visionary leader, and visionary PM and procurement components.
4. For the first time, strategic and tactical plans have been drawn up and will be used in the research test.

Conclusion

The GSA Heartland region is implementing an advanced and theoretically sound best value delivery process to transform the system from a price based to a best value environment. The major objectives include: minimization of management, direction, and control transactions, the transfer of risk and control to vendors who can minimize the risk, measurement of performance of the vendors and the GSA organization, and to measure an increase in performance and value of the services being delivered. A core group of visionaries are attempting to transform the organizational approach from one of management of personnel to a systems management, where performance measurements drive alignment of resources. This is a significant effort for a large federal organization that normally is management based and has difficulty in minimizing bureaucracy.

References

- Alsup, L. (2010, January). Branch Chief (Ret.), Design & Construction Division, Region 6. [Interview with researchers].
- Deming, E.W. (1982). *Out of the crisis*. Cambridge, MA: Massachusetts Institute of Technology.
- Kashiwagi, D. (2010). *Best value PIPS/PIRMS*. Performance Based Studies Research Group, Kashiwagi Solution Model Inc., Mesa, AZ.



- Kashiwagi, J., Sullivan, K., & Kashiwagi, D. (2009). *Risk management system implemented at the US Army medical command*, 7(3), 224-245.
- PBSRG. (2010). *Performance based studies research group internal research documentation*. Unpublished raw data. Arizona State University.
- Sullivan, K., Egbu, C., & Kashiwagi, D. (2005, February 8-10). Forcing contractors to improve with minimized management effort. *CIB W92 Construction Procurement: The Impact of Cultural Differences and Systems on Construction Performance*, 2, 683-691.
- Topi, J. (2010, January 9). Deputy Director Design & Construction Division, Region 6. [Interview with researchers]. Kansas City, KS.
- Kashiwagi, D.T., Kashiwagi, J., & Savicky, J. (2009). Industry structure: Misunderstood by industry and researchers. *Journal of Research*, VI(2), 59-76.



Panel #21 – Innovations in DoD Acquisition Policy

Thursday, May 13, 2010

1:45 p.m. –
3:15 p.m.

Chair: John T. Dillard, Senior Lecturer, Naval Postgraduate School

***Improving Defense Acquisition Processes with Evidence-based Analysis:
An Illustrative Case Using the DoD SBIR Program***

Toby Edison, Defense Acquisition University

Improving Defense Acquisition Decision-Making

William Fast, Defense Acquisition University

US Space Acquisition Policy: A Decline in Leadership

Barry Borst, National Reconnaissance Office, Shahram Sarkani and
Thomas Mazzuchi, George Washington University



Improving Defense Acquisition Processes with Evidence-based Analysis: An Illustrative Case Using the DoD SBIR Program

Toby Edison—Maj Toby Edison is a Professor of Program Management for Defense Acquisition University. Edison is a doctoral fellow at the RAND Graduate School. His thesis is an evaluation of the DoD SBIR program. While a program manager for Space Radar and Joint STARS he initiated several acquisition innovations: the GMTI Community of Practice and the GMTI Characterization lab. The GMTI Community of Practice brings together stakeholders to improve GMTI capabilities without an established program of record. This community provided guidance on GMTI for the ISR Task Force. The GMTI Characterization lab demonstrated and fielded several novel GMTI applications and CONOPs.

Maj Toby Edison
Professor of Program Management, Defense Acquisition University
222 N. Sepulveda Blvd, Suite 1220.
El Segundo, CA 90245-5659
Toby.edison@dau.mil
Ph: 310-606-5924 Fax: 310-606-5925

Abstract

This paper proposes and demonstrates that experimental and quasi-experimental program evaluation methods can be applied to some parts of the defense acquisition system to provide evidence of program effectiveness. The specific example presented is a quasi-experimental evaluation of the Department of Defense Small Business Innovation Research program. Quasi-experimental methods are a set of program evaluation techniques that allow researchers to approximate the results of an experimental study, such as a randomized controlled trial, without performing the experiment. The paper performs a quasi-experimental evaluation of the DoD SBIR program, which provides evidence that the program is effective at transitioning SBIR-funded technologies into other DoD programs. This demonstration that quasi-experimental methods can be used to evaluate certain aspects of the DoD acquisition system provides policy analysts with new tools to meet Congressional requirements for acquisition system evaluation. The paper recommends that more quasi-experimental studies be conducted and actual experimental studies be executed. These methods can help the DoD overcome the well-documented deficiency in evaluating the effectiveness of its acquisition systems. The Office of Management and Budget, the Government Accountability Office and the House Armed Services Committee unanimously agree that the DoD does not objectively measure the performance of its acquisition system.

Motivational Quotes

Findings.-The Congress finds that-

- (1) waste and inefficiency in Federal programs undermine the confidence of the American people in the Government and reduces the Federal Government's ability to address adequately vital public needs;*
- (2) Federal managers are seriously disadvantaged in their efforts to improve program efficiency and effectiveness because of insufficient articulation of program goals and inadequate information on program performance; and*



(3) Congressional policymaking, spending decisions, and program oversight are seriously handicapped by insufficient attention to program performance and results.

-- Introduction to the Government Performance Results Act of 1993

(Sec. 5403) Directs each federal agency required to participate in the SBIR or STTR program to:

(1) develop metrics evaluating the effectiveness and benefit of the program which are scientifically based, reflect the agency's mission, and include factors relating to the economic impact of the programs;

(2) conduct an annual evaluation of their SBIR and STTR programs using such metrics; and

(3) report each evaluation's results to the Administrator and the small business committees.

- Public Law 111-84, signed by President Obama on October 28, 2009, (authorizes National Defense for FY2010, and specifically authorizes the DoD SBIR/STTR Programs through September 30, 2010)

The Panel began with the question of how well the defense acquisition system is doing in delivering value to the warfighter and the taxpayer. For the most part, the Panel found that there is currently no objective way to answer this question. For most categories of acquisition, only anecdotal information exists about instances where the system either performed well, or poorly. Even where real performance metrics currently exist, they do not fully address the question. The Panel strongly believes that the defense acquisition system should have a performance management structure in place that allows the Department's senior leaders to identify and correct problems in the system, and reinforce and reward success.

- House Armed Services Committee Panel on Defense Acquisition Reform Findings and Recommendations, March 23, 2010

Introduction

Evaluating the effectiveness of any government program is difficult. Data on the program's output is often hard to obtain, selection into the program is usually not random and few programs are structured to facilitate the application of causal effects analysis. The Department of Defense (DoD) Small Business Innovation Research (SBIR) program is one such government program. Evaluating the effectiveness of the DoD SBIR program is required by Congress, who directs each federal agency to "develop metrics evaluating the effectiveness and benefit of the program which are scientifically based, reflect the agency's mission, and include factors relating to the economic impact of the programs." Despite this legal requirement and nearly 30 years of running the SBIR program, neither DoD administrators, nor policy analysts evaluating the program know whether the program is actually effective in supporting the DoD R&D mission by transitioning SBIR-funded technologies into DoD weapons systems. In their most recent assessments, the Government Accountability Office and the Office of Management and Budget, found that the effectiveness of the DoD SBIR program has not been demonstrated (GAO, 2005; OMB, 2005). The SBIR program is not alone in the DoD for its lack of evidence.

The indeterminate effectiveness of the relatively small SBIR program is just one case of the DoD generally not examining its acquisition processes. Congress finds that the Department of Defense acquisition system does routinely use objective methods to measure and improve its functions. Specifically, on March 23, 2010, the House Armed Services



Committee on Defense Acquisition Reform concluded that there is no objective way to determine “how well the defense acquisition system is doing in delivering value to the warfighter.” (HASC, 2010) Congress has officially required evidence-based policy administration by all Federal Agencies since 1993 through the *Government Performance and Results Act (GPRA)*. The GAO finds fault with the DoD’s implementation of the GPRA, finding serious flaws in the DoD’s Program Management business processes, which are responsible for managing DoD acquisition. Specifically the GAO cites, that the DoD’s plan to improve program management “lacked basic information, such as identifying specific business areas and key elements, such as goals, objectives, and performance measures.” (GAO, 2010) There is ample evidence that DoD’s measurement of its acquisition processes needs improvement. Unfortunately for many of the complex and unique acquisition processes that the DoD manages, instituting suitable performance measures has proved difficult. This paper shows that performance measurement tools exist for one small piece of the defense acquisition portfolio—the DoD SBIR program.

This paper proposes a methodology for measuring the performance of the DoD SBIR program that adapts quasi-experimental methods from the broader program evaluation literature. The paper begins with a description of the DoD SBIR program. It then describes the basics of the DoD SBIR program and examines two key biases in past DoD SBIR program evaluations that have confounded researchers: response bias and selection bias. The paper then documents strategies to mitigate these biases using quasi-experimental methods that have been used in other program evaluations. Next, the paper illustrates that a better evaluation of the DoD SBIR program is possible if better methods are applied to existing data. The paper then offers suggestions for strengthening evaluation of the SBIR program with better data collection methods and with randomization. With evidence that better evaluations of defense acquisition processes are possible, the paper concludes with suggestions for further evidence-based research.

Description of the DoD SBIR Program and Biases in Past Evaluations

Congress requires that all federal agencies with extramural R&D budgets in excess of \$100M, including the Department of Defense, set aside 2.5% of their R&D budget for the SBIR program. The broad purpose of the program is to provide contracts to qualifying small businesses to support each agency’s research mission, and to commercialize the funded technologies. In 2010, the SBIR program represents about 1% of the \$108B that the Department of Defense spends on procurement. Congress sets the emphasis of the program with the following four goals: 1) to stimulate technological innovation; 2) to use small businesses to meet federal R&D needs; 3) to foster participation by disadvantaged businesses; and 4) to increase private sector commercialization of federally funded research (OSADBU, 2007). Congress places more emphasis on the goal of increasing private sector commercialization.

The law also requires the participating federal agencies to structure their SBIR programs with three-phases, with specific funding ceilings for each phase. Phase I funds up to \$100K for a 6-month feasibility study competitively awarded to firms. Phase II is the principal R&D phase, which awards up to \$750K over 18 months to the most promising Phase I submissions. Phase III is the commercialization phase, which is the period when firms sell their mature technologies to interested parties—often DoD prime contractors or program offices. No pre-allocated SBIR program funds support Phase III commercialization; however, if a topic reaches Phase III, the firm can be awarded a contract for that technology immediately, without competition. The design of the SBIR Phases is intended to transition



the most promising technologies from the thousands of ideas of the participating small contractors into fielded technologies.

Within the constraints of the program, Congress offers freedom for the agencies to manage the SBIR program to fit the R&D strategies of the participating agencies, which are important to understand in order to evaluate the program. Each agency has many noteworthy organizational innovations for managing a large dollar R&D program without explicit overhead that is required to award contracts and grants in relatively small dollar amounts. The 2008 DoD annual report to Congress on the SBIR program highlights some of these challenges. In 2008 the DoD solicited proposals for nearly 1,000 topics, for which they processed over 12,000 proposals, ultimately awarding about two SBIR contracts per topic. In order to manage this administrative workload, the DoD manages the process online—publishing two or three SBIR solicitations a year online, requiring proposers to register with the DoD SBIR program with their unique federal contractor identification number and to submit their proposals online. These online contract management tools will be shown later to be invaluable for measuring the program effectiveness.

As highlighted in this paper's introductory quote from the 2009 re-authorization of the SBIR program, Congress requires the program administrators to develop metrics on the program's effectiveness. The DoD has created a metric called the Commercialization Achievement Index. This index is not deemed sufficient to measure the program's effectiveness (OMB, 2005). More broadly than the specific DoD program, across all federal SBIR programs, since its inception the effectiveness of the program to increase commercialization has never been evaluated (GAO, 2005). Among the specific reasons the GAO cites are lack of an agreed-upon measure of effectiveness for commercialization and lack of reliable data on the program. Published evaluations of the SBIR program typically suffer from two common issues identified in the broader literature on program evaluation: selection bias and response bias.

The key aspects of past DoD SBIR program evaluations that are presumed to cause bias are the fact that evaluations must be performed after the fact of selection and with self-reported survey data. Response bias affects program evaluations that rely on surveys because it is presumed that program participants over-report the output resulting from the program. Participants have an incentive to attribute more benefit from program participation in a survey so that the program will continue to receive funding and the participants continue to receive the benefits of the program. Selection bias is the presumption that program administrators are not selecting program participants at random. Specifically, selection bias invalidated after-the-fact evaluations because it is assumed that more capable participants are selected at a higher rate and that these firms, in the absence of the program, are more productive. In the case of the DoD SBIR program analyzed in this paper, winning firms were bigger, older, and more experienced defense contractors and as a result had more non-SBIR defense contracts before and after winning a SBIR award.

An ideal experiment of the SBIR program would randomly assign SBIR program treatment on a population of firms qualifying for the SBIR program and see if the treated firms have more future defense contracts than untreated firms. Such an experiment has not been conducted, which motivates the example in this paper, estimating the treatment effect for winning a DoD SBIR award with after-the-fact evaluation methods and non-survey data.

Strategies to Mitigate Biases

To perform a better effectiveness evaluation on the DoD SBIR program this paper builds a data set based on 2003 SBIR applications. To control for response bias, the applications were matched to the defense contract database rather than to survey data.



The analysis uses after-the-fact quasi-experimental models to control for selection bias, which have been shown to approximate the results of a randomized controlled trial under certain assumptions.

The program evaluation literature documents that the least biased program evaluations rely on a neutral source of outcome data (i.e. not reported by administrators or participants), have pre-treatment and post treatment observations, contain many characteristics of the participants and collect data on the treated population and a representative control population. The data set created for this analysis uses defense contract award data as the outcome of interest. The contract award data are an output of the defense accounting process represented by the DD Form 350, which documents and publishes every contract award greater than \$25K. The DoD identifies each contract awardee with a unique contractor identification number, which can be linked electronically to other databases the DoD maintains. This paper links to the DoD's Central Contractor Registry (CCR) and the DoD SBIR program's database of SBIR applications to capture firm characteristics in the database. The characteristics of each firm are important to after-the-fact program evaluations, because researchers can explain some of the variation in program effectiveness by correlating program outcomes with firm characteristics. For example, larger firms might win more defense contracts dollars simply because they have the capacity to take on more DoD-funded work, regardless of whether they won a SBIR award. The DoD SBIR program's database of SBIR applications captured information on all firms that applied for the DoD SBIR program by year of application and identified the firm's proposal that won an award. These pieces of information enabled the identification of a treatment population which applied for and won a SBIR award in a given year and a control population of firms that applied for but did not win an award. Creating a comparable control group with distinguishing characteristics is the crucial ingredient identified by program evaluation literature to controlling selection bias.

To control for selection bias the current program evaluation literature suggests using doubly robust estimation (DRE) methods to estimate the relationship between winning a SBIR award and future defense contract dollars. As the name implies, researchers use two methods to estimate a treatment effect. The first method prescribed is propensity score matching (PSM), which uses the observable covariates of the firms to create balanced treatment and control population. The second method prescribed is to perform a statistical estimation of the treatment effect that uses the characteristics of the firms to explain variation in future defense contracts (usually a regression with controls model). By combining two different estimation strategies, researchers have two chances to build the correct model. According to DRE theory, this approach will estimate a consistent treatment effect even if only one of the models is correct. The characteristic of double robustness is achieved in after-the-fact program evaluations when the estimation from the PSM model and the statistical model are consistent in magnitude and significance. Under ideal conditions and with enough descriptive data, by applying these methods, a better estimate of the treatment effect from winning a SBIR award on future defense contract dollars is possible.

A Naïve Estimate of SBIR Treatment

In order to show why using a balanced treatment and control population is better than using raw data, this paper begins with a naïve estimate of the DoD SBIR program's treatment effect. Researchers with a treatment and control group typically estimate a treatment effect with a differences in differences estimate. The first difference is calculated by subtracting the outcome observed before treatment and after treatment for each group.



The second difference is equal to the difference in treatment between treated and non-treated observations.

A differences in differences is not the same as a typical program evaluation report based on a survey. A survey based estimate can only report the average raw output data on the treated group. For example, the National Academies of Sciences reports the average raw survey response to estimate sales generated by SBIR funded research to be \$1.3M per SBIR project (Wessner, 2007). This average survey response is not a differences in differences because it does not compare the results to non-treated observations. Because the dataset created for this paper identifies winners and losers, it can be used to estimate a naïve differences in differences. Naïve means that that selection bias is not controlled.

The dataset used for this estimation is based on the entire population of DoD SBIR applicants in 2003 obtained from the Department of Defense SBIR administrative website. From the population of 2003 applications, a subset of 1460 firms who also applied in 2004, and who had a contractor identification number in the Central Contractor Registry, was identified as the population of interest. The DoD SBIR administrative database identifies 687 of these firms as winning a 2003 SBIR contract, with 773 applying for, but not winning, in 2003. These 1460 firms were matched with their contractor identification numbers to the form DD350 database maintained by the Department of Defense Directorate for Information Operations and Reports. The DD350 contains all contract actions greater than \$25K organized by year and by individual contractor identification number.

Using the SBIR application dataset, the first difference between average total non-SBIR defense contract dollars won in 2004 minus the 2003 total $\Delta 04-03$ is \$650K for the average winner and \$203K for the average loser (see Table 3). The second difference, the average treatment difference between winners and losers, is \$447K. This naïve treatment effect is assumed to be affected by selection bias.

Table 1. Naïve Differences in Differences

Group/Year	2003	2004	$\Delta 04-03$
Winners	1,430	2,081	650
Losers	456	659	203
$\Delta W-L$	974	1,422	\$447K

The effect of selection bias is presumably the cause of the SBIR winners having on average of \$974K more in contracts than losers did in 2003, and \$1.4M more in contracts in 2004. Because winners have more contracts to start out with, and firms with more past contracts will probably win more future contracts before and after winning in 2003, it is impossible to isolate the effect of winning the SBIR award in 2003. To improve on this naïve estimate, more advanced statistical techniques are needed.

Evidence of a SBIR Treatment Effect

The naïve treatment effect estimate can be improved by using the characteristics of firms to explain some of their variation in treatment. The characteristics are used two ways to control variation. The first method to control variation using firm characteristics is to use an algorithm to balance the characteristics of the treatment and control populations. The balancing algorithm will discard outlying observations so that the treatment and control populations will be theoretically identical to a randomized controlled trial population. The second method to produce a better estimate of treatment effect using firm characteristics is to use the firm characteristics to explain variation in the outcome. For example by using a



pre-treatment observation of defense contracts before a firm wins a SBIR contract, some of the variation in the post-treatment contract award amounts can be explained.

Applying these two methods to the dataset build for this paper can better estimate a treatment effect for the DoD SBIR program. This research method is described by Ho, Imai, King, and Stuart (2007) as doubly robust estimation. Double robust estimation protocols prescribe balancing populations and then using statistical methods to estimate the treatment effect. Analysis in Ho, Imai, King and Stuart (2007) shows consistency between the results of RCT studies analyzed with DRE methods. Their analysis supports the conclusion that estimates of causal treatment effects can be produced by DRE methods if researchers properly balance the treatment and control groups or researcher apply the correct statistical model. Their analysis based on thousands of different population balancing assumptions and statistical models with data from randomized controlled trials supports the conclusion that if the average treatment effect estimated with balanced treatment and control groups is consistent with the estimated treatment effect from another statistical model (such as a regression model) then the DRE estimate can be considered a causal estimate.

The model demonstrated estimates the future average increase in non-SBIR defense contracts for firms winning a 2003 DoD SBIR award. The key parameter of interest is the correlation between winning a 2003 SBIR award and non-SBIR defense contracts in 2004. The control variables are total non-SBIR contracts in 2002, total SBIR contracts in 2002, the firms' first contract year, the number of employees in 2003, whether the firm won a defense contract as a sub contractor in 2003, the number of topics submitted in 2003, and the total number of past Phase I or II awards.

The populations are balanced using the Coarsened Exact Matching protocols described by Iacus, King, and Porro (2008). The balanced population retains 534 firms that won in 2003 and 681 losing firms for a 83% post-matching retention rate. As an example of the improvement in post-matching balance, the raw population had a difference in 2002 non-SBIR contracts of \$925K, the matched population, \$58K.

The doubly robust estimation model estimates a \$147K treatment effect, with confidence level of greater than 99%. Based on this estimate, there is empirical support that the SBIR program increases defense contracts in 2004 for firms winning SBIR contracts in 2003.

The estimation that the DoD SBIR program does significantly increase non-SBIR defense contracts one year after award might be missing delayed effects two or three years after award. A three year commercialization time horizon is supported by surveys on the self-reported commercialization outcomes related to the SBIR program by the National Academies of Science (Wessner, 2007) and contract award analysis by RAND (Held, 2006), both of which find that the majority of commercialization activity occurs three years after a SBIR award. A doubly robust estimation is used to estimate several treatment effects for the non-SBIR DoD contracts won by firms in 2005 and 2006 who also won a 2003 DoD SBIR award. The doubly robust estimated treatment effect for the 2005 non-SBIR contract dollar difference is \$106K; the 2006 difference is \$130K. Both estimates are statistically significant at the greater than 99% confidence level. These estimations of a lagged treatment effect support a conclusion that for the average firm, winning a DoD SBIR award puts a company on a sustained path towards winning more future DoD contract dollars than had they not won.

Winning a DoD SBIR award appears to put winning firms on a path of higher non-SBIR defense contract award dollars. Figure 1 illustrates that for the period between 2004



and 2006 firms that applied for and won a 2003 SBIR contract won an average of \$370K more defense contracts than a matching set of firms who applied for but did not win a 2003 DoD SBIR award. The DoD SBIR program appears to be effective at increasing commercialization of SBIR funded technologies through defense contracts.

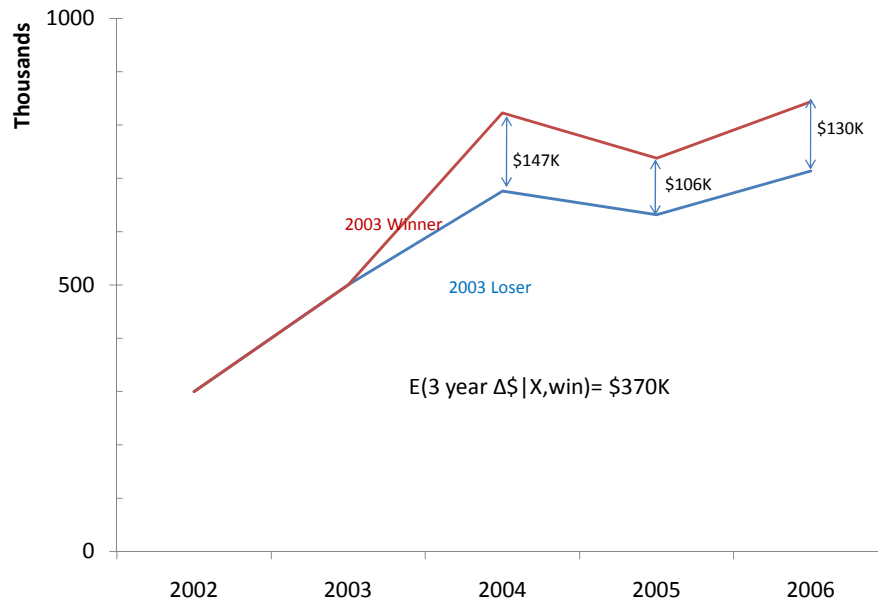


Figure 1. Three-year Estimated Treatment Effect of Winning a 2003 SBIR Award

The Department of Defense explicitly acknowledges that access to new technology and a strong industrial base are crucial to United States national security (OSD, 2010). The evidence presented in this paper suggests that the DoD SBIR program may be both providing access to new technologies and broadening the industrial base by transitioning new technologies developed by small businesses into defense programs through defense contracts. The evidence that firms winning SBIR contracts increase their future sales to the DoD at a higher rate than had they not won supports the belief that the DoD SBIR program contributes to the DoD mission. Prior to this analysis, the DoD emphasized without proof that they used the DoD SBIR program to support mission oriented research needs rather than to increase private sector commercialization. With proof that the commercialization path from SBIR funded R&D into standard defense prime contracts may be enhanced, the DoD can fulfill their GPRA requirement to demonstrate the effectiveness of their administration of the program and support the DoD preference against private sector commercialization.

This evidence can also provide a positive feedback loop for potential small business participants and program offices on the fence as to whether the program is worth their efforts. Higher quality potential contractors might be motivated to apply. Defense acquisition managers might be motivated to put more effort into developing SBIR topics and managing the technology transition process.

How to Improve DoD SBIR Program Evaluation

This analysis is motivated by a literature review of the SBIR program, which contains numerous government reports, policies and regulations requiring better evaluations of the DoD SBIR program. Most of the policy responses to the need for better evaluation, such as the DoD-developed Commercialization Achievement Index, and the surveys conducted by the GAO and National Academies of Science, fall short of actually providing data for better evaluation because the data collected is incomplete, presumably subject to response bias and does not collect data on treated and untreated populations. By using the already-existing defense contract database, this paper shows that there exists a data source free from self-reported survey response bias to evaluate the program. Additionally, by using econometric methods to control for selection bias, this paper provides policy makers with one example that it is possible to evaluate one key aspect of the program. The policy recommendations on how to improve evaluation will increase the number of studies on the program, allow researchers to explore more evidence of SBIR research output, and improve the policy recommendations of the program evaluations. This paper motivates three possible policy implementations the DoD can use to improve the evaluation of the DoD SBIR program. The first is to make the DoD SBIR administrative data accessible to more researchers. The second would be to build automated links to the applying SBIR firms to other innovation proxies—most specifically, the US Patent database, the iEdison database, and technical publication databases. Finally, to more conclusively evaluate the DoD SBIR program, some form of Randomized Control Trials will need to be implemented, and the enormous number of topics and applicants makes the DoD SBIR a good candidate to implement RCT's to evaluate the program.

Evaluation Recommendation 1: Make SBIR Administrative Data Available to Researchers

The first recommendation to improve evaluation of the program, making administrative data more accessible to researchers, is a low cost, easily implementable policy change with potential for significant payback. As already documented in the review of SBIR evaluations, one of the consistent themes of all past SBIR program evaluations is the lack of reliable, consistent data and the resulting lack of conclusive studies about the program's effectiveness. Additionally, the broader literature on R&D evaluations in general suffers from the same problems: lack of reliable data and a resultant dearth of conclusive evaluations on R&D programs. Opening the wealth of already-existing data collected by the DoD SBIR program to policy analysts would be an enormous step towards improving collective knowledge about how effective R&D subsidy programs really are. One specific example of data that is available to program administrators but not to program evaluators is the proposal evaluation scores used to award SBIR contracts. If these scores were made available to researchers, then researchers could use those scores to better match firms in propensity score models or to control for variation in outcome. Importantly, since the DoD SBIR program is probably already collecting this information for administrative purposes in electronic formats and making the data to available to administrators via the internet, the cost to make the data accessible to R&D policy researchers would be minimal. The payback for making this data available to research policy analysts that have spent decades trying to determine the efficacy of R&D policies with nearly zero reliable data is potentially significant. Policy makers could have more fact-based studies to improve policy to meet the spirit and intent of the Government Performance and Results Act.



Evaluation Recommendation 2: Link SBIR Funding to More Innovation Proxy Data Sources

The second policy recommendation to improve the evaluation of the DoD SBIR program is to enable automated matching of SBIR administrative data to other sources of innovation output data such as patent data, innovation tracking databases, sales data, venture capital funding, or technical publication data. Per US law, any SBIR participant is mandated to report to the government the details of any inventions or patents generated from the program. Unfortunately, the reporting is often decentralized, and the data collected is not easily linked to the actual source of funding. There are certainly more research outputs than just increased DoD sales tracked through the defense contracting database that could be used to measure the impact of the DoD SBIR program. Examples of potentially useful data sources are the US Patent and Trademark Database, technical paper databases, databases of firms such as COMPUSTAT, HOOVERS or DUNS, venture capital tracking databases, initial public offering databases, merger databases, or Internal Revenue Service data. Currently automated linking of SBIR participant data to another data source is not possible because not all of the databases can be linked using contractor identification numbers or DUNS numbers. The lack of a common standard firm identifier leaves researchers with the option of trying to match research inputs to output based on firm names, which contain tremendous variation in spelling within and across databases. The SBIR program could require firms to include their DUNS number in the already-required government interest statements for patents generated by SBIR funds. For matching technical publications, the SBIR program could require firms to report SBIR-generated technical publications with full citations in future application packages. Since SBIR application packages are submitted electronically, the government can begin to understand the impact of the SBIR program on the body of technical knowledge through patent disclosure analysis and technical publication analysis.

The most expedient link to establish might be the link between SBIR funding and the interagency Edison (iEdison) database maintained by the National Institutes of Health. This database was created to fulfill the statutory requirement for federally funded researchers to report inventions and patents developed with Federal funds. Currently it collects data from some, but not all, DoD research organizations. DoD SBIR policies could be modified to require winning firms to report inventions and patents through this database, and to require the inclusion of the funding contract number and the correct contractor identification number.

A final suggestion to improve tracking of SBIR output activity would be to require proposing firms to submit their tax identifier number to conclusively link SBIR funding to actual growth in revenue. Since all firms winning SBIR awards must be US companies, this policy intervention would cover the entire population of awardees. Moreover, since the IRS reports on income are legally required to be accurate and are subject to the possibility of auditing, the validity of the sales and revenue data will be substantially more accurate than the data self-reported in surveys. Another strength of this source of data would be that the study population could be expanded beyond the non-representative sample of survey respondents to include potentially all SBIR applicants.

The strengthening of the links between DoD SBIR program data sources and data sources on innovation proxies will greatly improve the quality and quantity of analyses possible on the program. If any of these policy recommendations improve evaluating the link between innovation subsidies to innovation output, a new era of R&D policy evaluation can begin and better R&D policies can be created.



Evaluation Recommendation 3: Implement Limited Randomized Control Trials for Improved Evaluations

The final suggestion for improving evaluation of the SBIR program is to continue to apply and refine research methods proven to mitigate biases, including using randomized controlled trials. The *Government Performance and Results Act* requires all agencies to strive towards evidence based policy implementation. The gold standard research method to provide conclusive evidence of program effectiveness would be to conduct a randomized control trial by randomizing some aspects of the contract awards. Of all the R&D subsidy and small business programs and the program evaluations reviewed for this paper, the SBIR program might be the most conducive to incorporating randomization to improve evaluation.

One practical suggestion to implement an RCT would be to select a subset of some of the topic awards with a random process. Since each topic receives around 15 applications, the suggestion would be to identify the five highest rated applications, randomly select the winner from those five applications, and track the relative performance of the firms that received the award and those who did not. There is a possibility that this type of experiment could be double blind because the firms would never know if they received the award due to random assignment and the program managers actually managing the SBIR contract could be kept blind to the actual award decision. The DoD SBIR program is an ideal candidate for incorporating some aspect of an RCT to evaluate the program. There are hundreds of topics each year, thousands of applicants, the research budget is by its very nature discretionary (not on a programs-critical path, or vital for national security), and the firms can be tracked over time.

In lieu of the opportunity to perform an RCT, researchers should continue to apply the propensity score and doubly robust estimation methods to SBIR administrative data. These after-the-fact estimation protocols could be improved if the actual evaluation scores were made available to researchers. If the evaluation scores were made available, researchers could use the scores to better match firms with balancing algorithms. Researchers could use the proposal evaluation scores in regression models to explain more variation in the outcomes of interest.

Current best practices in developmental economics have adopted RCT's (Rodrik & Rosenzweig, 2009). The focus of developmental economics—on improving the lives of the citizens of poor nations through interventions such as micro-financing, distributing anti-mosquito nets, improving immunizations and improving potable water supplies—by its nature makes it a much humbler and moderately funded field than national R&D policy analysis. Rodrik & Rosenzweig (2009) note that in the field of development economics:

- Randomized controlled trials (RCTs), in which randomly-selected subpopulations are selected for an intervention and then outcomes are compared across the treated and untreated populations, have been used to evaluate the causal effects of specific programs (e.g., cash transfers, subsidies to medical inputs), delivery mechanisms (e.g., kinds of financial products), and, less pervasively, to obtain evidence on fundamental behavioral assumptions that underlie models used to justify policy – e.g., adverse selection.

If policy administrators can adopt RCT methods to determine the best way to deliver developmental economics policy interventions, then the better-funded, higher-profile field of R&D policy analysis should be able to muster the resources and institutional will necessary



to implement limited RCT studies to better understand the efficacy of the \$1B+ DoD SBIR program.

Policy makers should seriously consider incorporating randomization into the DoD SBIR program to improve the evaluation of the program and to demonstrate how to build evaluation tools into other government programs.

Conclusion on How to Improve SBIR Program Evaluation

These three suggestions could help revolutionize the way the SBIR program is evaluated and offer a wider variety of answers to the policy questions. With more data available, better links to research output and actual experimental results, the artifacts of the DoD SBIR program that actually work best can be understood, refined and applied as best practices across the DoD and Federal government. With better analyses, policy makers can use facts to craft and administer better policies. This paper has provided a small sample of the research possible if evaluation data and tools are improved. If any form of these recommendations is adopted, the DoD SBIR program can be better evaluated.

Suggestions for Further Evidence-based Acquisition Policy Analysis

- The program evaluation tools demonstrated in this paper highlights that it is possible to evaluate the effectiveness of some aspects of the defense acquisition systems. The after-the-fact tools demonstrated in this paper and the suggestion to implement randomized controlled trials can be applied to other areas of the defense acquisition system to provide policy makers evidence of how well policy changes perform. Specifically there are policy changes enacted by the *Weapon System Acquisition Reform Act* (P.L.111) and the *National Defense Authorization Act for FY 2009* (P.L. 110-417) that are worthy of consideration for evaluation with experimental and quasi-experimental methods. Some examples of the policy recommendations that might be suited for experimental analysis are as follows: the emphasis on competition, the requirement for prototyping, the implementation of earned value management, and the increase in the number of acquisition professionals.
- For example, estimating the effectiveness of maintaining competition throughout the acquisition lifecycle could be part of a randomized trial or could be analyzed using quasi-experimental methods. For an RCT, policy makers could randomly pick which current program would be required to implement competition in technology development, prototyping, and production. Analysts could estimate the effect of competition by measuring the difference in cost changes and schedule delays on the programs with and without competition. If randomization of competition requirements is infeasible, after-the-fact analyses could estimate the effect of competition on cost and schedule. The evaluator could use the characteristics of the different programs (weapon type, joint program, service of program office, year of program initiation), along with an identifier on whether they had competition or not to build treatment and control groups and to explain other variations in program outcomes.

Conclusion

Congress is re-emphasizing its direction to the DoD to improve the evaluation methodologies for the defense acquisition system. This paper highlights that for some



aspects of the defense acquisition system quasi-experimental methods can be applied and do provide evidence to estimate program efficacy. This paper recommends that DoD policy makers build more experimental and quasi-experimental links into the current DoD SBIR program to improve the evidence available to acquisition policy makers. Based on this demonstration, policy makers should consider broadening the application of these methods beyond the SBIR program to acquisition system aspects that can be analyzed with experimental and quasi-experimental models.

References

- DIOR. (n.d.). *DoD personnel & procurement statistics*. Retrieved from <http://siadapp.dmdc.osd.mil/>
- GAO. (2010, March 25). *DoD approach to business transformation*. Retrieved March 23, 2010, from http://www.gao.gov/highrisk/risks/DoD-management/business_transformation.php
- GAO. (2005). Observations on the small business innovation research program: statement for the record of Anu K. Mittal, Director Natural Resources and Environment Team. Washington DC: Author.
- GAO. (2009). Reports on the Government Performance and Results Act. Washington DC: Author.
- House Armed Services Committee Panel on Defense Acquisition Reform (HASC). (2010). *House Armed Services Committee Panel on defense acquisition reform, findings and recommendations*. Washington DC: Author.
- Held, B.E. (2006). Evaluation and recommendations for improvement of the Department of Defense small business innovation research program. RAND: Santa Monica.
- Ho, D., Imai, K., King, G., & Stuart, E.A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15(3), 199-236.
- Iacus, S., King, G., & Porro, G. (2008). *CEM: A matching method for observational data in the social sciences*. Retrieved November 15, 2009, from www.agrocampus-ouest.fr/math/useR-2009/.../pdf/Porro+Iacus+King.pdf
- OMB. (2005). OMB program assessment rating tool (part): DoD Small Business Innovation Research Program. Washington, DC: Author.
- Office of Small and Disadvantaged Business Utilization (OSADBU). (2007). *DoD annual report summary*. Washington, DC: Author.
- Office of Small and Disadvantaged Business Utilization (OSADBU). (2010, February 26). *DoD SBIR annual report*. Retrieved February 26, 2010, from <http://www.DoDsbir.com>
- OSD. (2010). *Quadrennial defense review report*. Washington DC: Author.
- Rodrik, D., & Rosenzweig, M. (2009). Development policy and development economics: An introduction*. In D. Rodrik & D.R. Rosenzweig (Eds.), *Handbook of development economics* (Vol. 5). North-Holland.
- Wessner, C.W. (Ed.). (2007). An assessment of the small business innovation research program at the Department of Defense. Washington, DC: National Academies Press.
- US Congress. (2009). *Weapon system acquisition reform act* (P.L.111). Washington, DC: US GPO.
- US Congress. (2009). *National defense authorization act for FY 2009* (P.L. 110-417). Washington, DC: US GPO.



Improving Defense Acquisition Decision-Making

William Fast—COL William R. Fast, USA (Ret.), facilitates financial and program management courses at the Defense Acquisition University. He also writes and speaks on various management topics and provides consultation services to defense acquisition programs. From 2001-2004, COL Fast worked in the Office of the Assistant Secretary of the Army (Acquisition, Logistics and Technology), where he was responsible for ensuring that weapons systems acquisition programs were properly reflected in the Planning, Programming, Budgeting, and Execution system.

COL William R. Fast, USA (Ret.)
Defense Acquisition University
9820 Belvoir Road
Fort Belvoir, VA, 22060-5565
Work Phone: 703-805-2107
Home Phone: 703-922-5169
Fax: 703-805-3186
e-mail: william.fast@dau.mil

Abstract

This research investigates evidence and tests the hypothesis that the linkages between the defense acquisition management system, the requirements process, and the budgeting system are not sufficiently defined to enable the success of acquisition programs. These disconnects contribute to weapons systems cost overruns, schedule delays, and performance problems, and are exacerbated by the ever-changing global security environment and rapid pace of technological advancement. Through historical research, qualitative and quantitative analyses, and a comprehensive review of current policies and procedures, this research illuminates these areas of disconnect and proposes specific recommendations to fix them.

Keywords: Acquisition, Budgeting, Decision-making, Programming, Requirements

The primary purpose of this research was to investigate how well the Defense Acquisition Management System interfaces with the requirements and budgeting systems of the Department of Defense (DoD). The United States of America possesses the finest weapons systems in the world. However, the same cannot be said for the systems that enable the Pentagon to acquire those weapons systems. Cost overruns, schedule delays, and operational test failures testify to numerous severed connections among the acquisition management, requirements, and budgeting systems (commonly referred to as the three decision support systems). The ever-changing global security environment and the rapid pace of technological change only serve to exacerbate these problems.

For the Pentagon to earn a reputation for excellence in acquiring weapons systems, these decision support systems must operate with far better coordination and demonstrate that they can procure the right equipment, within reasonable timeframes, and at affordable prices. This research began with an investigation into the intricacies of the acquisition management, requirements, and budgeting systems. Next, interactions between these three decision support systems were illuminated to uncover areas of misalignment and disconnect. Recent initiatives to correct these problems were also identified. Finally, solutions to resolve these disconnects were enumerated.



Background

A January 2006 report of the Defense Acquisition Performance Assessment (DAPA) described the three decision support systems as:

a highly complex mechanism that is fragmented in its operation. Further, the findings we developed indicated that differences in the theory and practice of acquisition, divergent values among the acquisition community, and changes in the security environment have driven the requirements, acquisition, and budget processes further apart, and have inserted significant instability into the acquisition system. In theory, new weapons systems are delivered as the result of the integrated actions of the three interdependent processes whose operations are held together by the significant efforts of the organizations, workforce, and the industrial partnerships that manage them. In practice, however, these processes and practitioners often operate independent of one another. Uncoordinated changes in each of the processes often cause unintended negative consequences that magnify the effects of disruptions in any one area.⁵⁴ (DAPA, 2006, pp. 4-5)

1. This problem has not been fixed. Writing in the January/February 2009 issue of *Foreign Affairs*, Secretary of Defense Robert M. Gates (2009) called for a reassessment of priorities within the Department of Defense:

The defining principle of the Pentagon's new National Defense Strategy is balance. The United States cannot expect to eliminate national security risks through higher defense budgets, to do everything and buy everything. The Department of Defense must set priorities and consider inescapable tradeoffs and opportunity costs.

The strategy strives for balance in three areas: between trying to prevail in current conflicts and preparing for other contingencies, between institutionalizing capabilities such as counterinsurgency and foreign military assistance and maintaining the United States' existing conventional and strategic technological edge against other military forces, and between retaining those cultural traits that have made the U.S. armed forces successful and shedding those that hamper their ability to do what needs to be done. (p. 28)

How Gates will achieve this rebalancing of priorities is the essence of this research.



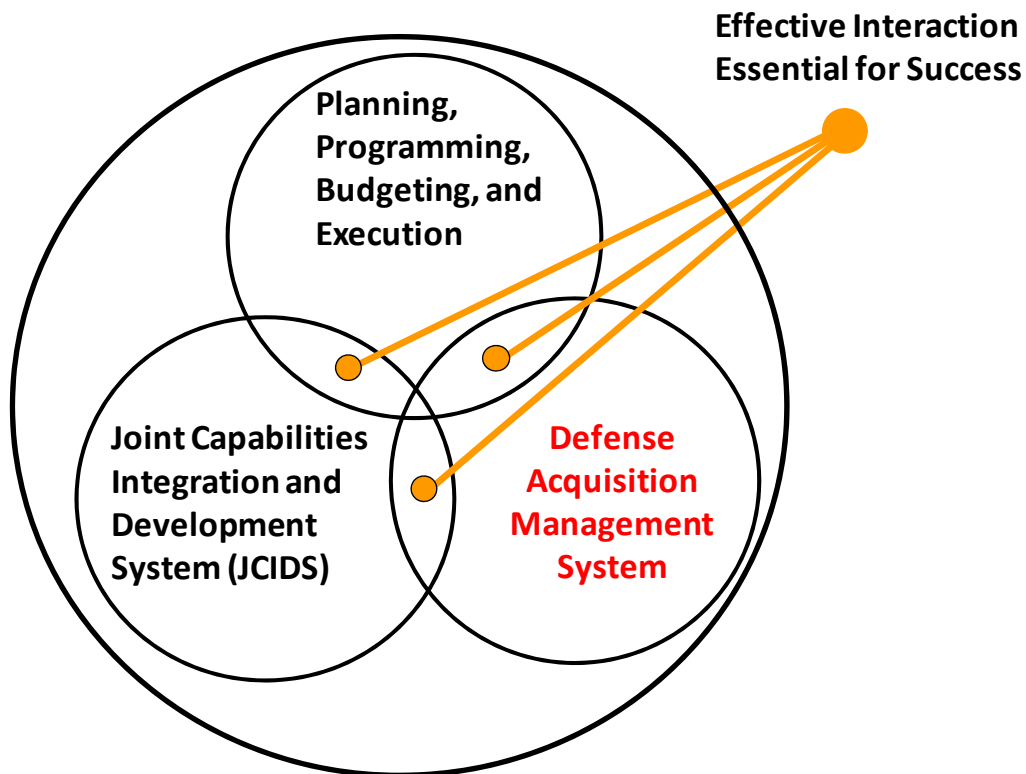


Figure 1. DoD Decision Support Systems

Figure 1 highlights the areas of interaction between the Defense Acquisition Management System, the Joint Capabilities Integration and Development System (JCIDS), and the Planning, Programming, Budgeting, and Execution (PPBE) system. Coordinated management decisions at these interfaces are essential for the success of any acquisition program. Thus, this research began by seeking to understand the reasons why these three decision support systems were first established and how acquisition programs are affected by the decisions made within and between these systems today.

Defense Acquisition Management System (DAMS): Stratified Decision-Making

Decision-making in today's Defense Acquisition Management System (DAMS) can be traced to 1986. The late David Packard, then president of Hewlett-Packard, was selected by Ronald Reagan to lead the President's Blue Ribbon Commission on Defense Management. Better known as the Packard Commission, its interim report of April 1986 recommended the appointment of both DoD-level and Service Acquisition Executives (SAEs). The SAEs would appoint Program Executive Officers (PEOs) under their authority that would be responsible for a manageable number of acquisition programs and project managers. By design, the chain of authority from the project manager, through the PEO, to the SAE was short. The basic premise was that defense acquisition needed to be streamlined to run in the same manner as a commercial venture (Butrica, 2001, pp. 212-213).

Another feature of the acquisition management system is that it classifies programs for higher levels of oversight based upon expected development or production expenditures. An Acquisition Category I (ACAT I) Major Defense Acquisition Program (MDAP), requiring oversight by the Defense Acquisition Executive (DAE) or DoD Component Acquisition Executive (CAE), if so delegated, is a program that is expected to require in excess of \$365 million of Research, Development, Test, and Evaluation (RDT&E) funds and/or \$2.19 billion of procurement funds (in fiscal year 2000 constant dollars) (DoD, 2008d, encl. 3, p. 33).

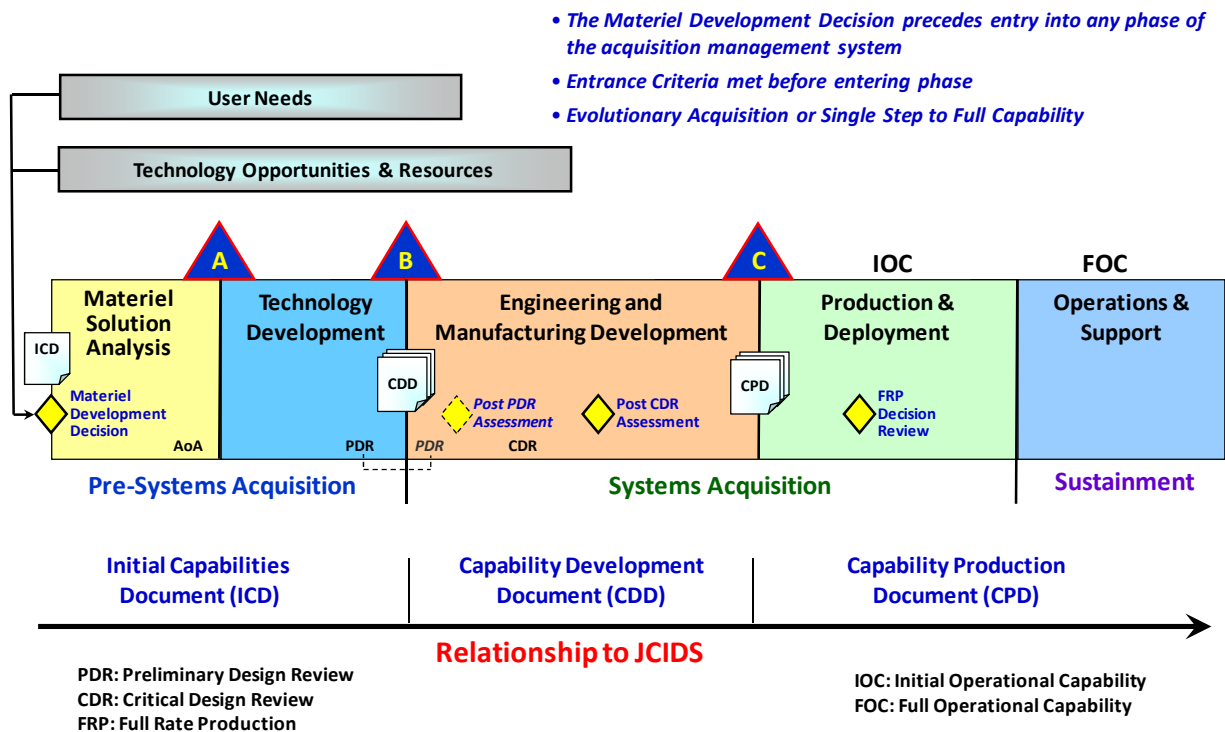


Figure 2. The Defense Acquisition Management System

Unlike the PPBE process that is *calendar-driven* or the JCIDS which is *needs-driven*, the acquisition management system is *event-driven*. All acquisition programs are managed through a series of sequential phases and milestone reviews (Figure 2). To successfully move from one phase to the next, a program must have demonstrated or completed the program-specific exit criteria for the current phase and must also have met the statutory and regulatory entrance criteria for the next phase. The appointed Milestone Decision Authority (MDA) makes the “go/no-go” decision based on the evidence presented at the milestone review.

The effect of having a higher-level decision-maker for MDAPs is that 31% of the department’s programmed Research, Development, and Acquisition (RDA) funds are under the authority of one decision-maker—the Under Secretary of Defense for Acquisition, Technology and Logistics, who is the designated DAE. Yet, the remaining 69% of

programmed RDA funds are under the control of the Services and Defense Agencies.⁵⁵ The total number of decision-makers with MDA for lower-priority acquisition programs is over 40.⁵⁶

In addition, analysis of acquisition decision memoranda (ADMs) documenting the decisions of the DAE for MDAPs reveals that 36% of the ADM contained language with impact on the requirements decision-making process, and 66% of the ADMs contained actions affecting decisions in the budgeting process.⁵⁷ Obviously, decisions made on the more numerous lower acquisition category programs also ripple into the requirements and budgeting processes at higher rates.

JCIDS: Centralizing the Validation of Capability Documents to Ensure “Jointness”

Historically, the military Services have had their own systems for the approval of weapons systems requirements. However, in 1976 the Office of Federal Procurement Policy published *Circular A-109* that required a Mission Area Analysis to determine the need for a particular weapons system (OMB, 1976). In compliance with *A-109*, the Services were required to perform this analysis and prepare a mission needs statement to document the need at the front end of the acquisition process (Fox & Field, 1988, p. 46). Eventually, to ensure that requirements were not duplicated between the Services and to prompt interoperability and joint operations, the Joint Staff got involved. In the early 1990s, they required the Services to adopt a single document format for the Operational Requirements Document (ORD). In 2003, the Joint Capabilities Integration and Development System (JCIDS) process was created to identify the capabilities and associated operational performance criteria required by the joint warfighter. JCIDS also supports the statutory responsibility of the Joint Requirements Oversight Council (JROC) to validate joint warfighting requirements.

⁵⁵ In Future Years Defense Program 2008-2013 (FYDP 2008-2013), the total obligation authority for RDT&E and Procurement was \$1,154 billion. By virtue of the fact that the Under Secretary of Defense for Acquisition, Technology and Logistics is the MDA for MDAPs, the OSD has control over acquisition decisions totaling \$362 billion, or about 31% of the total obligation authority in FYDP 08-13. On the other hand, the Services make decisions on about \$792 billion, or about 69% of the total obligation authority for RDT&E and procurement in FYDP 08-13 (DoD, 2008c, Table 1-9, p. 13; DAMIR, n.d., MDAP/MAIS Selected Acquisition Report query, FYDP 08-13).

⁵⁶ Each Service and Defense Agency has an Acquisition Executive (AE) with MDA. In addition, all PEOs have MDA. The total number of PEOs is 35 (Army-11; Navy-13; Air Force-11). (Organizational charts of Army, Navy, and Air Force AEs, retrieved November 14, 2009, from <https://www.alt.army.mil/portal/page/portal/oasaalt>; <https://acquisition.navy.mil/rda/content/view/full/4539>; and <http://ww3.safaq.hq.af.mil/organizations/index.asp>).

⁵⁷ The Defense Acquisition Management System (DAMS) uses ADMs as records of the decision made by the AE. For the purposes of this research, ADMs for the following weapons systems were reviewed: Expeditionary Fighting Vehicle (6 ADMs); Future Combat System (10 ADMs); Global Hawk (12 ADMs); Joint Strike Fighter (13 ADMs); and Littoral Combat Ship (3 ADMs). In total, 44 ADMs were reviewed. Of these, 36% (16 ADMs) contained actions that would require involvement of the JCIDS. In addition, 66% (29 ADMs) contained actions that would impact upon the PPBE process (ADM, n.d.).



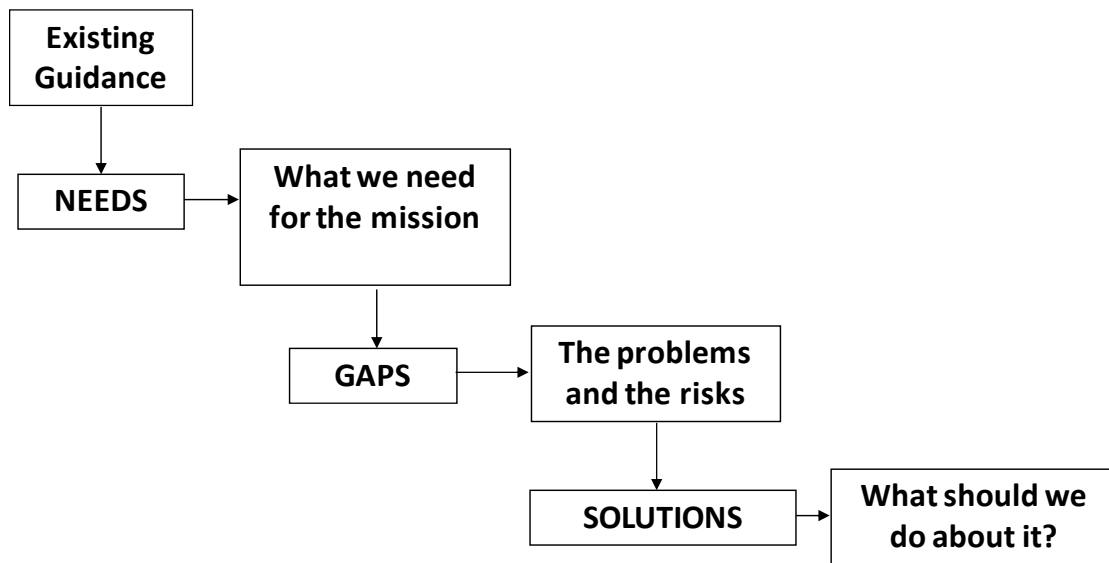


Figure 3. Capabilities-based Assessment

Fundamental to JCIDS is Capabilities-based Assessment (CBA) (Figure 3). Unlike the more predictable threats of the Cold War that the Pentagon could anticipate and prepare for, threats today emerge on a daily basis, and are often asymmetrical to our existing capabilities. CBA seeks to find solutions to these emerging threats by changing Doctrine, Organization, Training, Material, Leadership and Education, Personnel, and Facilities (DOTMLPF) (CJCS, 2009, p. GL-3). The CBA process produces initial capability, capability development, and capability production documents (ICD, CDD, and CPD). These documents guide the technology development, engineering and manufacturing development, and production and deployment phases of the acquisition framework, respectively (Figure 2).

Chairman of the Joint Chiefs of Staff Instruction (CJCSI) 3170.01G is explicit regarding how JCIDS interfaces with the two other decision support systems:

The JCIDS process supports the acquisition process by identifying and assessing capability needs and associated performance criteria to be used as a basis for acquiring the right capabilities, including the right systems. These capability needs then serve as the basis for the development and production of systems to fill those needs. Additionally, it provides the PPBE process with affordability advice by assessing the development and production life-cycle cost. (CJCS, 2009, pp. A-1, A-2)

An approved ICD summarizes the CBA process, describes the capability gaps, and identifies potential solutions. The ICD is taken to a Materiel Development Decision (MDD), where it is reviewed and validated in order to start the acquisition process. A favorable MDD leads into the Materiel Solution Analysis phase, which is prior to Milestone A. In this phase, an Analysis of Alternatives (AoA) is prepared, based upon the broad type of materiel solution preferred in the ICD (i.e., information system, evolutionary development of an existing capability, or a transformational approach) (CJCS, 2009, p. A-3). Each alternative has an associated life-cycle cost that gives insight into the affordability of the program and provides linkage to the budgeting process. So, it is important to note that the information in

the ICD drives the AoA process. The ICD also informs the technology development strategy, the test and evaluation strategy, and the systems engineering plan—all key documents for guiding the technology development phase prior to program initiation at Milestone B.

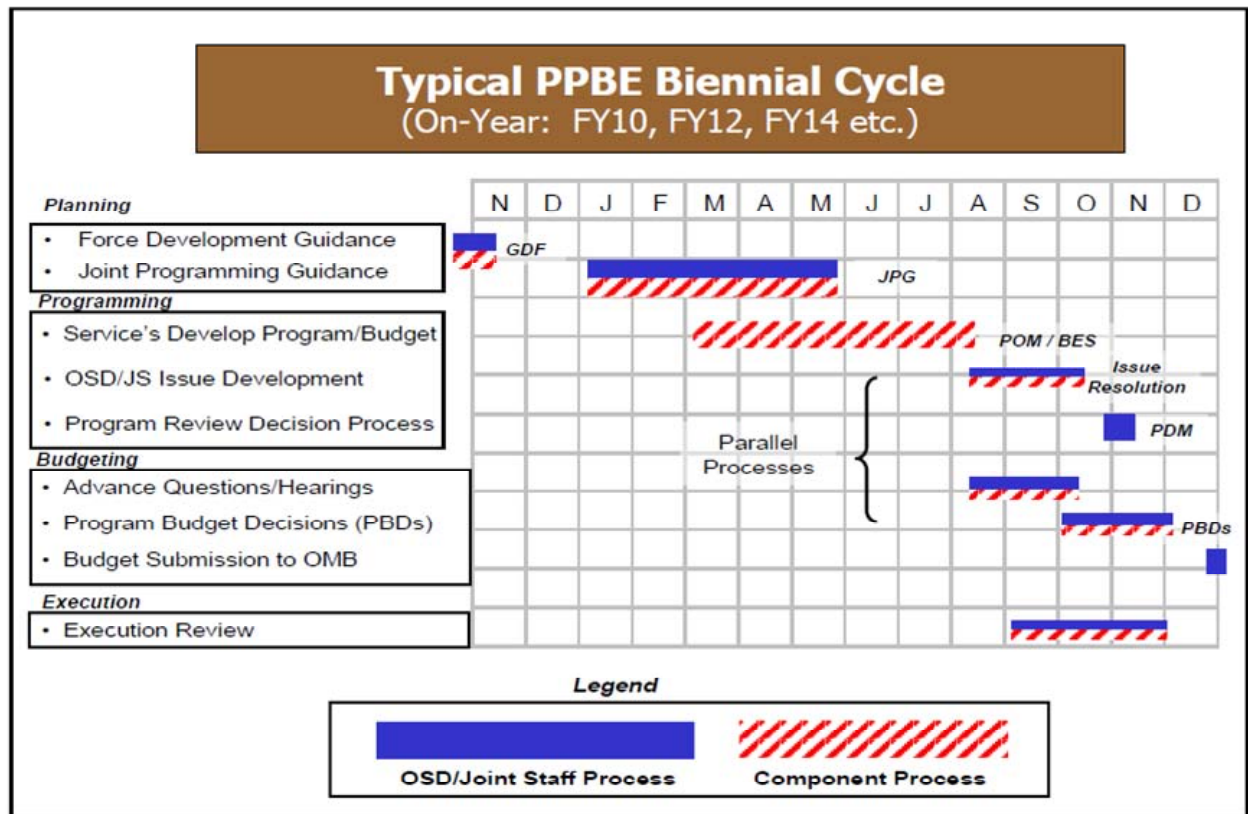


Figure 4. Planning, Programming, Budgeting, and Execution

PPBE: Then and Now

In the spring of 2008, the American Society of Military Comptrollers (ASMC) surveyed 575 members of the defense financial management community about the PPBE process (Figure 4). Agreement was almost universal that PPBE was the best method to link performance and budgeting, “and a strong sentiment to fully implement the system as designed” (ASMC & Grant Thornton, 2008, p. 1). So, just what was PPBE originally designed to do? And, has the DoD implemented PPBE in a way that allows it to do what it was designed to do? To find answers to these questions, one must go back to the beginnings of PPBE (then PPBS), during the era of Defense Secretary Robert S. McNamara.

In 1961, President Kennedy’s initial instructions to McNamara were “to determine what forces were required and to procure and support them as economically as possible” (McNamara, 1964, p. 14). Developed by cost analysts at the RAND Corporation during the 1950s, program budgeting was just what the Pentagon needed to link budget inputs to capability outcomes and to centralize long-range planning and financial decision-making under the civilian Secretary of Defense. The system was originally called the Planning, Programming, and Budgeting System (PPBS), and its fundamental purpose was to unify annual budgets and nonfinancial longer range planning. In the age of the nuclear bomb, the

task of long-range planning was to calculate the needed effects or outputs that had to be produced by military forces and weapons systems in order to prevail. Budgeted funds for these military forces and weapons systems came from the funding appropriations for military personnel, research and development, procurement, and operations and maintenance. Yet, budgets are resource inputs. Moreover, because of the long development cycles for modern weapons systems, annual budgeting was not a useful planning tool. The key for McNamara, and the objective of PPBS, was to link the planning outputs to the appropriated funds inputs through the construct of defined program elements within a 5-year force structure and financial program (Novick, 1962, p. 2).

As originally envisioned, planning within the PPBS was to be a comparative analysis of the projected costs and effectiveness of feasible alternatives. The example used by David Novick, one of the developers of program budgeting, is the comparison of the merits of buying more Polaris submarines versus Minuteman missile squadrons. Both systems could deliver nuclear warheads. The comparison between the two alternatives involved the methodical examination of the cost estimates for manpower, equipment, and facilities, and the expected military benefits (capability outcomes) derived from the systems (Novick, 1962, p. 6). Today, comparatively little analysis to this level of detail takes place in the planning phase of PPBE. Up until 2006, planning was simply an effort to turn the *National Security Strategy (NSS)*, *National Defense Strategy (NDS)*, *National Military Strategy (NMS)*, and the *Quadrennial Defense Review (QDR)* into guidance from which the Services could develop their Program Objective Memoranda (POMs). Such a shallow planning effort resulted in guidance that was not specific enough, in terms of priorities and quantities, for the programming of adequate resources for weapons systems acquisitions. Here is but one of many examples.

The *National Security Strategy* (Clinton, 2000) was silent on the role of the military in finding and taking the fight to terrorists. While the document discusses the need for the military to help deter terrorism and respond in retribution to terrorist attacks, the mission of finding and destroying terrorist organizations is not mentioned. Thus, the FYDP for fiscal years 2002-2007, prepared by the Pentagon in fiscal year 2000, lacked a vision for the weapons systems and equipment necessary to prosecute an offensive global war on terror (Paparone, 2008, p. 157).⁵⁸ As the world changes at an unprecedented pace, casting a meaningful strategic vision becomes more and more problematic. Without meaningful strategic vision, the acquisition management system may continue to acquire programs that will no longer be needed—and may fail to start programs that will be needed. The Obama

⁵⁸ COL Christopher R. Paparone, USA (Ret.), makes an argument that the *Joint Vision 2020*, published in June 2000, focused on defensive force protection from terrorists, not on the use of military forces to combat terrorism in an offensive way, which was the case after September 11, 2001. While the *Joint Vision 2020* was not a PPBE document, per se, his point is applicable. Combating terrorists offensively is not seen in the *National Security Strategy* prior to 9/11. This is not the only failure on the part of past presidential administrations in providing meaningful strategic priorities. The 2002 *National Security Strategy (NSS)* failed to envision the invasion of Iraq on March 20, 2003, the fall of Baghdad, and the associated requirements for nation building that were thrust onto the military. The 2004 *National Military Strategy (NMS)* failed to envision the need for massive humanitarian aid in the wake of the Indian Ocean tsunami of December 24, 2004, and the associated requirements that the military would need for logistical support across the shores of devastated islands and coastal regions. Similarly, the 2005 *National Defense Strategy (NDS)* failed to envision that North Korea would test fire missiles over the Sea of Japan on July 4, 2006, and subsequently explode a nuclear device in the mountains on October 9, 2006. The 2005 *NDS* makes no mention of our nation's need to acquire an integrated missile defense capability.



Administration has yet to set clear national security priorities. As a result, the Pentagon began in early 2009 the planning phase for fiscal years 2012-2017 without the benefit of an NSS. Clearly, no one knows what the future will hold. However, planning for a future we cannot see and attempting to bring that illusion to the future fight, with all the associated weapons systems acquisition requirements, is clearly folly if not patently dangerous. Yet, this is the current planning process upon which the Pentagon justifies and builds its 6-year defense program.

The *Government Performance and Results Act (GPRA)* of 1993 requires that each government agency establish a results-oriented management approach to strategically allocate resources on the basis of performance (US Congress, 1993). In assessing the implementation of *GPRA*, the Government Accountability Office has criticized the DoD for not establishing goals or timelines for accountability and for the measurement of progress toward implementation. The DoD implemented a risk management framework in its strategic plan—the 2001 *QDR* report (GAO, 2005, p. 8). However, it was not until 2003 that the DoD adopted the balanced scorecard approach to implement risk management. The GAO criticized the DoD for not integrating this framework with other decision-making support processes. Specifically, the GAO said that to be effective, risk-based and results-oriented management approaches have to be integrated into the usual cycle of agency decision-making. The GAO presumed that without this level of integration, a mismatch between programs and budgets would continue, and a proportional rather than strategic allocation of resources would go to the Services.⁵⁹ In addition, the Congress would not have insight as to the risks and trade-offs made during the Pentagon's investment decision-making (GAO, 2005, p. 5).

Chartered to examine how the DoD develops, resources, and provides joint capabilities, the Joint Capabilities Study Team (also called the Aldridge Study) reported these findings to Defense Secretary Rumsfeld in 2004: "Services dominate the current requirements process...; Service planning does not consider the full range of solutions

⁵⁹ The Government Accountability Office says that even though the DoD has adopted a risk management planning framework and balanced scorecard approach to programming for outcomes, the percentage of total obligation authority in the FYDP, by Service, has remained relatively unchanged. The GAO provided the following figures in its report *Defense Management: Additional Actions Needed to Enhance DoD's Risk-based Approach for Making Resource Decisions* (GAO-06-13):

Table 1. Military Service and Defense-wide Percentage of the 2005 and 2006 Future Years Defense Programs
(GAO, 2005, p. 16)

	2005 Percentage of FYDP	2006 Percentage of FYDP	Percentage Change by Department
Department of the Army	24.23	24.63	0.40
Department of the Navy	29.75	29.47	-0.28
Department of the Air Force	29.80	29.82	0.02
Defense-wide	16.22	16.08	-0.14
Total	100.00	100.00	

available to meet joint warfighting needs...; and, the resourcing function focuses senior leadership effort on fixing problems at the end of the process, rather than being involved early in the planning process.” They also found that programming guidance exceeds available resources (DoD, 2004, p. iii). Others have also identified this programming guidance “gap” (Christie, 2008, p. 196; Church & Warner, 2009, p. 82; Johnson, 2003, p. 9).

The Aldridge Study proposed a four-step process: strategy, enhanced planning, resourcing, and execution and accountability. The strategy step involved the combatant commanders and answered the question: “What to do?” The enhanced planning and resourcing steps answered the question: “How to do it?” The execution and accountability step answered the question: “How well did we do?” Formal process review points for the Secretary of Defense were proposed after each of the four steps (DoD, 2004, p. v).

Many of the recommendations from the Aldridge Study were implemented. Most notably, the Enhanced Planning Process (EPP) was made a phase of the Strategic Planning Process, and the EPP is to be approved by the Secretary or Deputy Secretary of Defense. Moreover, the Joint Programming Guidance is to document the decisions resulting from the EPP phase (DoD, 2006, p. 2). The Director of Program Analysis and Evaluation (PA&E) already had responsibility as the Office of the Secretary of Defense (OSD) lead for coordinating the program review of the PPBE process. The only problem with this new assignment is that it appears to conflict with the responsibilities of the Under Secretary of Defense (Policy), who has overall responsibility for coordinating the PPBE planning phase (DoD, 2003, p. 5).

Another problem for PPBE is that developing and finally enacting the first year of the 6-year program takes a long time. The program (termed *Future Years Defense Program* or *FYDP*) is put together only once every 2 years, during even numbered years. For example, in calendar year 2010, the Pentagon will put together the 6-year program for fiscal years (FY) 2012 through 2017. However, the Services began working on their portions of that FY 2012-2017 program in the middle of calendar year 2009—more than 3 years before the first year funds for FY 2012 will be appropriated by the Congress. The next opportunity to make major changes to the program is in calendar year 2012 when the program for FY 2014-2019 will be accepted by the Pentagon. Changes to the program are possible during the odd numbered years. However, these changes are usually limited to necessary fact-of-life adjustments. New starts (or stops) are generally not considered in the odd numbered years. Thus, the programming phase of the PPBE process suffers from *false precision*. Even if the vision of the future were correctly identified in the planning phase, programming for weapons systems new starts can only be done every other year. Moreover, funds requested are for use more than three or more years hence. Inevitably, projections for weapons systems costs that far in advance of execution are bound to be flawed. Yet, the process demands precision, whether or not that precision has any meaning (McCaffery & Jones, 2005, p. 159).

As originally envisioned, Secretary McNamara expected to conduct a continuous review of the entire defense program. In other words, he expected to have an up-to-date 5-year force structure and financial program at all times. McNamara’s PPBS had a program change control system in which variations from approved cost estimates required advance authorization. Standard forms were established for research and development, investment, and operations—each relating to the key decision points in the life of a weapon system. The program change control system was first applied to 200 of the most important material systems. Milestone schedules were prepared for these systems, and actual progress was reported on a monthly basis, including the need for corrective action or revision to the



financial plan (Novick, 1962, pp. 7-10). Such is not the case today. The FYDP is open for changes only twice a year—in August when POMs (or changes to the previous POMs) are submitted by the Services to OSD, and at the end of the combined program and budget review once resource management decisions have been made and the defense budget is finalized for the Office of Management and Budget.

In his first year as Defense Secretary, McNamara was heavily involved in the cost-effectiveness and requirements studies of the planning phase of PPBS. Known as “McNamara’s 100 Trombones,” he assigned about 100 requirements projects to the Joint Chiefs of Staff, the Services, and various elements of OSD. These planning studies were truly participative in nature and required a significant time commitment from McNamara, but they resulted in detailed acquisition programming guidance for the Services. For example, in his first year, McNamara made decisions on the number of strategic missiles and bombers for the next decade. He also decided on the airlift and sealift needed to support contingency war plans and the most cost-effective way of replacing worn out ground equipment for the Army (Hitch, 1965, pp. 74-75).

Senior leader involvement in today’s PPBE process has typically been toward the end of the programming phase rather than in the earlier planning phase. This is not the optimum time for these senior leaders to enter the PPBE decision-making process. Moreover, failing to make the tough decisions up front in the planning phase only delays them into late in the programming phase (Johnson, 2003, pp. 10-11). Decisions become harder to make during the final stages of programming because less discretionary funding is available, and earlier decisions will need to be reconsidered. Such late decision-making on weapons systems acquisition terminations was typical in past PPBE cycles. However, as demonstrated by Defense Secretary Gates during the 2010 budget deliberations, he may get more involved up front and make these types of decisions early in the planning phase of PPBE.

Today, PPBE fiscal and programming guidance is usually late in arriving to the Services. While no directive or instruction establishes a date for issuance of fiscal/programming guidance, issuance dates for the past two PPBE cycles were March 14, 2008, for POM 10-15; and May 7, 2009, for POM 11-15. Fiscal guidance refers to the total obligation authority, by fiscal year, available to the Services. Fiscal/programming guidance is used by the Services to develop their POMs, or changes to the previous POM, which are usually due in August. They begin development of their POMs in the last few months of the prior year (October-December timeframe). While draft fiscal/programming guidance is often released earlier, final fiscal/programming guidance is usually issued too late to be useful. Today, fiscal/programming guidance is found in the “fiscally informed” Guidance for the Development of the Force (GDF) and the “fiscally constrained” Joint Programming Guidance (JPG) (Church & Warner, 2009, p. 84). The predecessor to the GDF was the Strategic Planning Guidance (SPG), and before the SPG, the Defense Planning Guidance (DPG). Originally envisioned to align strategy with investments, the GDF appears to have become a “wish list of programs and priorities for every constituency.” Feedback from the Services on the usefulness of the GDF and JPG is mixed. As indicated, both documents, but especially the JPG, are issued well after the Services have completed the development of their POMs and decisions have been made to fund or not fund various weapons systems programs (Church & Warner, 2009, pp. 81-82).

Understandably, and working at a disadvantage with unclear programming guidance, the Service POMs are invariably criticized for failing to comply with the GDF/JPG. In addition, the POMs are faulted for underestimating technology risks associated with



weapons systems investments (Christie, 2008, p. 212). As a result, the Services tend to *over program*, believing they can develop, produce, and place in operation many more programs than realistically possible (Christie, 2008, p. 196; Church & Warner, 2009, p. 82). In other words, their 6-year programs fail to consider the cost “tails” past the last year of the FYDP. This is particularly a problem with weapons systems production programs that build up to an unrealistically high “bow wave” of procurement funding beyond the FYDP that becomes unaffordable for the Service and the DoD.

Per *DoD Directive 7045.14*, the official linkage between the PPBE and acquisition management systems is achieved by *designated membership* on the Defense Systems Acquisition Review Council (now the Defense Acquisition Board (DAB)), the Defense Resources Board (now the Deputy’s Advisory Working Group (DAWG)), and the Senior Leader Review Group (SLRG); and the requirement to develop an acquisition strategy for all major systems (DoD, 1984, reissued 1987, p. 6). The DAB is chaired by the Under Secretary of Defense for Acquisition, Technology and Logistics, who is also a member of the SLRG and DAWG. The SLRG is chaired by the Secretary of Defense, and the DAWG is chaired by the Deputy Secretary of Defense, neither of whom sits on the DAB. In total, 11 senior leaders are members of both the DAB and the SLRG/DAWG.⁶⁰ The average tenure of the DAE is just 24 months.⁶¹ Most MDAPs have development cycles that exceed the tenure of four or even five DAEs. Therefore, the effectiveness of having senior leaders serve as the linkage between the resourcing and acquisition management systems might be questioned, given their enormous responsibilities and brief tenures serving as the DAE. Certainly, 11 senior leaders cannot be held responsible for coordinating the multitude of interactions between the acquisition and budgeting systems.

Recommendations

In 1979, the *Defense Resource Management Study (DRMS)* recommended to President Carter that the programming and budgeting phases of PPBE be combined into a single annual review. The *DRMS* also recommended that the time freed up by combining the two phases be used to “focus additional attention on the strategic and resource planning issues, including resolution of selected major issues prior to the program/budget review” (Rice, 1979, p. viii). This was the centerpiece of the *DRMS* proposal, and it was designed to open up a “broad planning window” that would include “an orchestrated OSD review and prioritization of the Defense Systems Acquisition Review Council-approved programs competing for segments of the planning wedge” (Rice, 1979, pp. 9, 16). These recommendations were not implemented. However, in 2003, Defense Secretary Rumsfeld did combine programming and budgeting phases, but not with the intention of freeing up time for better planning. Rather, Rumsfeld’s *Management Initiative Decision 913* specified

⁶⁰ The members of both the DAB and the SLRG/DAWG are as follows: Under Secretary of Defense for Acquisition, Technology and Logistics; Vice Chairman of the Joint Chiefs of Staff; Secretaries of the Military Departments; Under Secretary of Defense (Policy); Under Secretary of Defense (Comptroller); Under Secretary of Defense (Personnel and Readiness); Under Secretary of Defense (Intelligence); Assistant Secretary of Defense for Networks and Information Integration/DoD Chief Information Officer; Director, Cost Assessment and Program Evaluation (DoD, 2008b, encls. 3 & 4; DoD, 2009, p. 10.2.1).

⁶¹ From Richard Godwin, the first Under Secretary of Defense (Acquisition) until Ashton Carter, the current Under Secretary of Defense for Acquisition, Technology and Logistics, average tenure has been 24 months. To date, the shortest service was by Godwin, who served 12 months (September 1986–September 1987), and the longest service was by Jacques Gansler, who served 38 months (November 1997–January 2001) (Brown, 2005).



that the freed up time would be used for an execution review (i.e., the new “E” in PPBE) to “make assessments concerning current and previous resource allocations and whether the department achieved its planned performance goals” (DoD, 2003, p. 7; Church & Warner, 2009, p. 81; Dawe & Jones, 2005, p. 49; Jones & McCaffery, 2005, p. 90). The Pentagon has yet to institutionalize this execution review. A recent survey of 575 professionals in the defense finance and accounting community found that, due to the wartime supplemental funding for operations in Iraq and Afghanistan, emphasis on execution had not made the relationship between budget execution and performance more visible, nor had it provided the data needed to make more timely decisions to improve the PPBE process (ASMC & Grant Thornton, 2008, pp. 5-7). Perhaps, the “broad planning window” recommendation of the *DRMS* should again be considered, and this time implemented, to help resolve and clarify competing requirements and acquisition programs before the Services have to prepare their POMs.

In 2007, Capability Portfolio Management was introduced to the programming phase of PPBE. The official definition of Capability Portfolio Management is “the process of integrating, synchronizing, and coordinating DoD capabilities needs with current and planned DOTMLPF investments within a capability portfolio to better inform decision making and optimize defense resources” (DoD, 2008a, p. 8). The Capability Portfolio Management initiative seeks to place all current and proposed warfighting needs into logical, manageable functional categories. In an effort to minimize redundant capabilities, capability portfolios are joint, not Service-specific. Capability Portfolio Managers (CPMs) provide cross-Component alternatives and recommendations on current and future capability needs and investments. They are to work with the JROC and the JCIDS, and develop capability planning guidance for inclusion in the GDF. Therefore, CPMs can impact capability portfolio composition, weapons systems acquisition, and weapons systems sustainment choices. In retrospect, the job of the CPMs is similar to the system analysts of the McNamara era. The systems analysts prepared “cost-effectiveness studies” and “requirements studies” at the request of the Secretary of Defense and the Joint Chiefs of Staff (Hitch, 1965, pp. 73-75). However, the advice of current-day CPMs is officially sought only at the end of the programming phase of PPBE when they provide the DAWG with independent programmatic recommendations and cross-Component perspectives on planned and proposed capability investments (DoD, 2008a, p. 6). To have greatest influence, decision-makers need to formally tap into the advice of these CPMs about 9 to 12 months earlier, during the planning phase of the PPBE process.

The deliberate, evolutionary pace of the Cold War is long past. The challenges of an ever-changing global security environment and the rapid pace of technological advancement represent a national imperative for the Pentagon to seek out and cultivate breakthrough ideas in the development and employment of defense systems (Johnson, 2003, pp. 6-7). To meet these challenges, the PPBE planning phase should be revitalized and extended to allow time for brainstorming and germination of innovative ideas and for the analysis of the costs and effectiveness of various weapons systems alternatives.



Conclusions

As implemented today, the PPBE process is far different from the PPBS established by Secretary of Defense McNamara in 1961. Over the course of nearly 50 years, changes have severely de-emphasized decision-making in the planning phase. As a result, the Department has had to establish a separate requirements analysis and approval system. The concept behind today's JCIDS was actually part of McNamara's long-range planning to determine the most cost-effective capability outcomes. Likewise, in McNamara's management system, weapons systems development and production decisions, along with necessary funding adjustments, were made in real time, and at the same time as requirements decisions. Today, the linkage between PPBE and weapons systems decisions suffers from the timing disconnect between a calendar-driven budget and event-driven acquisition programs. To improve acquisition decision-making, the linkages between the requirements, budgeting, and acquisition decision-making systems must be reestablished. One solution is to reinvigorate the planning phase of PPBE and make the necessary decisions on weapons systems requirements, multi-year budgeting, and acquisition program continuation or termination, within the timeframe of that phase.

References

- Acquisition Decision Memoranda (ADM)*. (n.d.). Retrieved November 14, 2009, from <https://extranet.acq.osd.mil/dab/adm/index.html> (a secure web site accessible only from servers within the DoD by Common Access Card holders)
- American Society of Military Comptrollers (ASMC) & Grant Thornton LLP. (2008, May). *Supporting the warfighter: Linking budgets to results*. Alexandria, VA: Author.
- Brown, S. (Ed.). (2005). *Providing the means of war: Historical perspectives on defense acquisition, 1945-2000*. Washington, DC: US GPO.
- Butrica, A.J. (2001). An overview of acquisition, 1981-1990. In S. Brown (Ed.), *Providing the means of war: Historical perspectives on defense acquisition, 1945-2000* (pp. 199-223). Washington, DC: US GPO.
- Christie, T. (2008). *Long in coming, the acquisition train wreck is here*. In W.T. Wheeler (Ed.), *America's defense meltdown: Pentagon reform for President Obama and the new congress* (pp. 195-218). Washington, DC: Center for Defense Information.
- Church, A.T., & Warner, T. (2009). DoD planning, programming, budgeting, and execution system: A path toward improvement. *Joint Force Quarterly*, 53, 80-84.
- CJCS. (2004). *The national military strategy of the United States of America: A strategy for today; A vision for tomorrow*. Retrieved November 9, 2009, from <http://www.defenselink.mil/newsd/Mar2005/d20050318nms.pdf>
- CJCS. (2009). *Joint capabilities integration and development system* (Instruction 3170.01G). Washington, DC: Author.
- Clinton, W.J. (2000, December). *National security strategy*. Washington, DC: Author.
- Dawe, R.L., & Jones, L.R. (2005). Transformation of national defense business management: Current initiatives and future challenges. *International Public Management Review*, 6(2), 9-71.
- Defense Acquisition Management Information Retrieval (DAMIR)*. (n.d.). Retrieved November 13, 2009, from <http://www.acq.osd.mil/damir/>
- Defense Acquisition Performance Assessment (DAPA) Panel. (2006, January). *A report by the assessment panel of the defense acquisition performance assessment (DAPA) project for the Deputy Secretary of Defense*. Defense acquisition performance



- assessment (DAPA) report. Retrieved November 14, 2009, from <https://acc.dau.mil/CommunityBrowser.aspx?id=18554>
- DoD. (1984, administrative reissuance incorporating change 1, April 9, 1987). *Implementation of the planning, programming, and budgeting system (PPBS)* (DoDI 7045.7). Washington, DC: Author.
- DoD. (1984, certified current as of November 21, 2003). *The planning, programming, and budgeting system (PPBS)* (DoDD 7045.14). Washington, DC: Author.
- DoD. (2003). *Implementation of a 2-year planning, programming, budgeting, and execution process* (DoD MID 913). Washington, DC: Author.
- DoD. (2004). *Joint defense capabilities study: Improving DoD strategic planning, resourcing and execution to satisfy joint capabilities*. A report by the Joint Capabilities Study Team for the Secretary of Defense. Retrieved November 14, 2009, from <https://acc.dau.mil/CommunityBrowser.aspx?id=294616>
- DoD. (2006). *Director, program analysis and evaluation (PA&E)* (DoDD 5141.01). Washington, DC: Author.
- DoD. (2008a). *Capability portfolio management* (DoDD 7045.20). Washington, DC: Author.
- DoD. (2008b). *DoD senior governance council* (DoDD 5105.79). Washington, DC: Author.
- DoD. (2008c). *National defense budget estimates for FY 2009 budget (Green Book)*. Washington, DC: Office of the USD (Comptroller). Retrieved November 13, 2009, from http://comptroller.defense.gov/defbudget/fy2009/fy2009_greenbook.pdf
- DoD. (2008d) *Operation of the defense acquisition system* (DoDI 5000.02). Washington, DC: Author.
- DoD. (2009). *Interim defense acquisition guidebook (DAG)*. Washington, DC: Office of the USD (AT&L).
- Fox, J.R., & Field, J.L. (1988). *The defense management challenge: Weapons acquisition*. Boston, MA: Harvard Business School Press.
- GAO. (2005, November). *Defense management: Additional actions needed to enhance DoD's risk-based approach for making resource decisions* (GAO-06-13). Washington, DC: Author.
- Gates, R.M. (2009, January/February). A balanced strategy: Reprogramming the Pentagon for a new age. *Foreign Affairs*, 88(1), 28-32.
- Hitch, C.J. (1965). *Development and salient features of the programming system*. In S.A. Tucker (Ed.), *A modern design for defense decision: A McNamara-Hitch-Enthoven anthology* (pp. 64-85). Washington, DC: Industrial College of the Armed Forces.
- Johnson, S.E. (2003). A new PPBS process to advance transformation. *Defense Horizons*, 32, 1-6. Retrieved November 10, 2009, from http://www.ndu.edu/ctnsp/defense_horizons.htm
- Jones, L.R., & McCaffery, J.L. (2005). Defense acquisition and budgeting: Investigating the adequacy of linkage between systems. *International Public Management Review*, 6(2), 87-115.
- McCaffery, J.L., & Jones, L.R. (2005). Reform of program budgeting in the Department of Defense. *International Public Management Review*, 6(2), 141-176.
- McNamara, R.S. (1963, April). *Decision-making in the Department of Defense*. In S.A. Tucker (Ed.), *A modern design for defense decision: A McNamara-Hitch-Enthoven anthology* (pp. 20-26). Washington, DC: Industrial College of the Armed Forces.
- McNamara, R.S. (1964, April-June). *Managing the Department of Defense*. In S.A. Tucker (Ed.), *A modern design for defense decision: A McNamara-Hitch-Enthoven anthology* (pp. 11-19). Washington, DC: Industrial College of the Armed Forces.
- Novick, D. (1962, November). *Program budgeting: Long-range planning in the Department of Defense*. Santa Monica, CA: RAND Corporation.



- OMB. (1976). *Major system acquisitions*. Washington, DC: Office of Federal Procurement Policy.
- OSD. (2005, March). *The national defense strategy of the United States of America*. Retrieved November 9, 2009, from <http://www.defenselink.mil/news/Mar2005/d20050318nds1.pdf>
- Paparone, C. (2008). PPBE: A red or a blue pill? Can defense sensemakers really be rational in a hyperturbulent world? *Defense Acquisition Review Journal*, 16(2), 151-173.
- Rice, D.B. (1979). *Defense resource management study final report*. Washington, DC: US GPO.
- US Congress. (1993). *Government performance and results act (GPRA)* (Pub. L. No. 103-62). Washington, DC: US GPO.

Appendix

List of Abbreviations and Acronyms

ACAT	Acquisition Category
ADM	Acquisition Decision Memorandum or Acquisition Decision Memoranda
AE	Acquisition Executive
AoA	Analysis of Alternatives
ASMC	American Society of Military Comptrollers
BES	Budget Estimate Submission
CAE	Component Acquisition Executive
CBA	Capability-based Assessment
CDD	Capability Development Document
CDR	Critical Design Review
CJCS	Chairman of the Joint Chiefs of Staff
CJCSI	Chairman of the Joint Chiefs of Staff Instruction
CPM	Capability Portfolio Manager
DAB	Defense Acquisition Board
DAE	Defense Acquisition Executive
DAMS	Defense Acquisition Management System
DAPA	Defense Acquisition Performance Assessment
DAWG	Deputy's Advisory Working Group
DoD	Department of Defense
DoDD	Department of Defense Directive
DoDI	Department of Defense Instruction
DOTMLPF	Doctrine, Organization, Training, Material, Leadership and Education, Personnel, and Facilities
DPG	Defense Planning Guidance
DRMS	<i>Defense Resource Management Study</i>
encl.	enclosure
EPP	Enhanced Planning Process
FRP	Full Rate Production
FY	Fiscal Year
FYDP	Future Years Defense Program
GAO	Government Accountability Office
GDF	Guidance for the Development of the Force
GPRA	<i>Government Performance and Results Act</i>
ICD	Initial Capabilities Document



JCIDS	Joint Capabilities Integration and Development System
JPG	Joint Planning Guidance
JROC	Joint Requirements Oversight Council
JS	Joint Staff
MDA	Milestone Decision Authority
MDAP	Major Defense Acquisition Program
NDS	National Defense Strategy
NMS	National Military Strategy
NSS	<i>National Security Strategy</i>
OMB	Office of Management and Budget
OSD	Office of the Secretary of Defense
ORD	Operational Requirements Document
OSD	Office of the Secretary of Defense
PA&E	Program Analysis and Evaluation
PBD	Program Budget Decision
PDM	Program Decision Memorandum
PDR	Preliminary Design Review
PEO	Program Executive Office or Program Executive Officer
POM	Program Objective Memorandum or Program Objective Memoranda
PPBE	Planning, Programming, Budget, and Execution
PPBS	Planning, Programming, and Budgeting System
QDR	<i>Quadrennial Defense Review</i>
RDA	Research, Development, and Acquisition
RDT&E	Research, Development, Test, and Evaluation
SAE	Service Acquisition Executive
SLRG	Senior Leader Review Group



US Space Acquisition Policy: A Decline in Leadership

Barry “Jay” Borst—Jay Borst is a PhD student in systems engineering at the George Washington University. His current research interests include acquisition policy and innovation. He currently works at the National Reconnaissance Organization.

Shahram Sarkani—Professor Shahram Sarkani has been engaged in engineering research, technology development, and engineering education since 1980. The author of over one hundred technical publications and presentations, Professor Sarkani focuses his research on the study of how such natural forces such as earthquakes affect buildings, bridges, and other human-made structures. His main research interests are structural failure—with emphasis on structures subjected to dynamic and cyclic loadings—and the effects of earthquakes and other natural forces on infrastructure. In 1990, he founded, and still directs, GW's Center for Infrastructure Safety and Reliability, where he and his associates and students explore dynamic and cyclic loadings on structures to predict how long a structural system will last if the foundation under it undergoes occasional violent erratic motion. Their research is supported by the National Science Foundation, the US Navy, the National Aeronautics and Space Administration, and the United States Agency for International Development.

Professor Sarkani joined the faculty of the School of Engineering and Applied Science of George Washington University in 1986, after earning his PhD from Rice University. He served as chair of the Civil, Mechanical, and Environmental Engineering Department from 1994 to 1997. From 1997 to 2001, he was SEAS Interim Associate Dean for Research and Development. Since 2001, he has served as Faculty Adviser for Off-Campus Programs in the Department of Engineering Management and Systems Engineering.

Thomas Mazzuchi—Thomas Mazzuchi is a professor of Operations Research and Engineering Management at the George Washington University. His current research interests include reliability and risk analysis, Bayesian inference, quality control, Stochastic models of operations research, and time series analysis.

Abstract

System complexity is only one aspect affecting US space acquisition today. There is a large body of literature that suggests US space acquisition is over budget, behind schedule, and delivering underperforming systems. The GAO seems to attribute a number of factors to contributing to this situation. However, three primary factors include an over-reliance on immature technology, managing requirements to build the “grand design” and the health of the space industrial base. Addressing these factors will be critical so that the US can maintain its technology superiority and leadership in space. This is especially critical as countries such as Russia and China continue to mount significant challenges to our dominance in space. A loss of US leadership in space could very well translate into a loss of prosperity and national security.

In this new century, those who effectively utilize space will enjoy added prosperity and security and will hold a substantial advantage over those who do not. In order to increase knowledge, discovery, economic prosperity, and to enhance the national security, the United States must have robust, effective, and efficient space capabilities. (Office of Science, 2006)



Introduction

Since the onset of the Cold War, the United States has led the development of space for exploration, commercial uses, and national interests. In that time, space has managed to permeate almost every aspect of our society and culture. In the span of forty some years, space has enabled global communications, broadcasted various forms of entertainment, assisted in geolocating new deposits of natural resources such as gas and oil, predicted the weather, provided information on our adversaries, and facilitated the conduct of successful military operations.

Despite our success, the national security of the United States faces an increasing risk of threat as a number of factors mount a significant challenge to our leadership in space. The Government Accountability Office (GAO) and others observe that a majority of these challenges may lie with our acquisition practices, procedures and policies. Their reviews suggest that many of our developmental national space systems are over budget, significantly behind schedule and woefully inadequate in terms of expected user performance. Lt. Gen. Michael Hamel, former commander of Air Force Space and Missile Systems Center, was pointed when he stated, “Nothing threatens US military superiority in space more than a loss of ability to develop, field and sustain our space systems” (Hamel, 2009).

Noting the challenges confronting the US space acquisition program, this paper seeks to accomplish three objectives. First, is to briefly explain the importance of US leadership in space by describing not only its historical development but also its contribution to the prosperity and national security of the United States. Second, it seeks to explain a small subset of the current acquisition challenges the US is facing, to include: dependence on immature technology, requirements, and the state of space industrial base. Finally, in conjunction with those challenges, this paper offers some modest recommendations that the acquisition community might employ to overcome these challenges and ensure US leadership in space.

The Space Development Imperative

Shortly after World War II, the US became engaged in the Cold War against its former ally, the Soviet Union. From a strategic perspective, the US recognized that, first, it needed a capable and credible nuclear deterrent against the Soviet Union and that, second, it needed to understand Soviet military developments and intentions. The development of intercontinental ballistic missiles (ICBMs) solved the US’s need to deliver nuclear warheads half way around the world. However, on the second issue, it still needed a means of obtaining information from denied areas. The shoot down of Gary Powers and his U-2 aircraft only accentuated the need for space systems capable of accessing denied areas. In other words, space development was the imperative.

Many organizations had aspirations for and were vying for control of space, and it wasn’t until the 1960s that the Air Force and the National Reconnaissance Office (NRO) would become the principal developers of military space power for the nation. During that time, there were rapid advances in communications, weather, navigation, missile warning, and intelligence surveillance (Hamel, 2009). These advances continued for the remainder of the Cold War. However, these space assets were typically only employed in a strategic and operational role. As the Cold War came to a close, and as Iraq flexed its military power in the Persian Gulf War, the perspective of using space systems as only strategic and operational assets quickly changed.



On August 2, 1990, Iraqi forces invaded the Kingdom of Kuwait, and the Iraqi government almost immediately annexed Kuwait as the nineteenth province of Iraq. From the early moments of the first Persian Gulf War, it became apparent that satellites were not only a force multiplier on the strategic and operation level but also on the tactical level. For the first time, satellites connected military forces, sensors, and decision-makers across the battlespace; collected data on operationally relevant conditions, reconnoiter, survey and target hostile forces; and enabled precision, synchronization and command and control of forces in the field (Hamel, 2009). As a result, Russia and China watched with disbelief as the American military easily dismembered and neutralized a Soviet-trained and -equipped military in a matter of days.

Since the close of the Persian Gulf War, the US military has increasingly relied on satellites to conduct and facilitate military operations. Additionally, the nature of warfare has evolved as adversaries recognize and attempt to negate the advantage that space systems provide for their American users. As a result, US military forces have asked for increasingly complex space systems in greater numbers and on faster timelines to answer the challenge posed by adversaries. This, in part, has contributed to the current state of US space acquisition.

The Acquisition Challenge

Without significant improvements in the leadership and management of national security space programs, U.S. space preeminence will erode 'to the extent that space ceases to provide a competitive national security advantage.' (Chaplain, 2009)

Recent studies conducted by the GAO suggest that major weapon system acquisition is a serious cause of concern for government leaders. For example, the GAO found that out of the DoD's portfolio of 96 programs, "42% are higher than originally estimated and the average delay in delivering initial capabilities has increased to 22 months" (Chaplain, 2009). Table 1 displays how costs have risen and schedules expanded for "big acquisition" from 2003 to 2008, using fiscal year 2009 dollars.

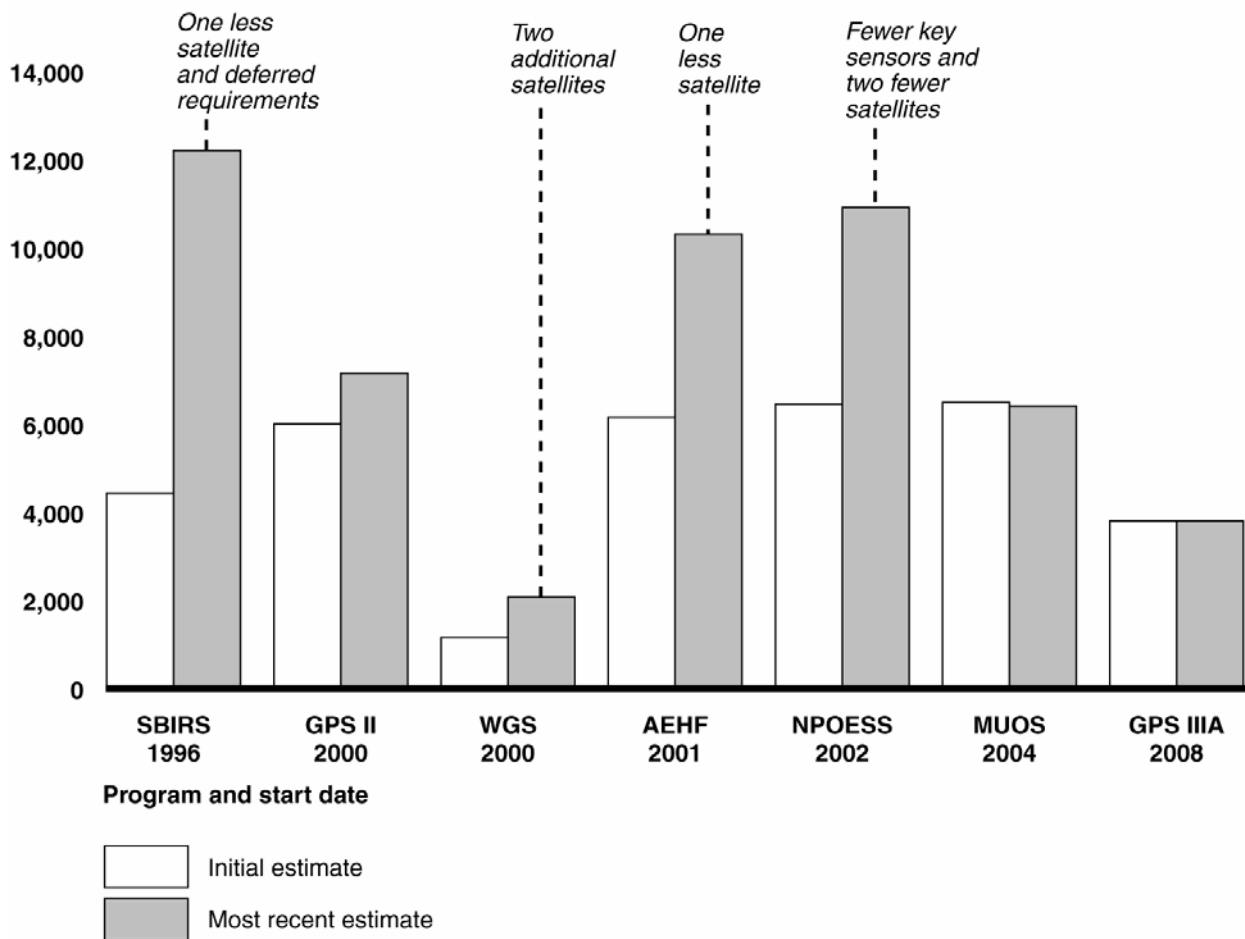


Table 1. Analysis of the DoD's Major Defense Acquisition Program Portfolios
(Francis, 2009)

Fiscal Year			
	2003	2007	2008
Portfolio size			
Number of Programs	77	95	96
Total Planned Commitments	\$1.2 Trillion	\$1.6 Trillion	\$1.6 Trillion
Commitments Outstanding	\$742.2 Billion	\$875.2 Billion	\$786.3 Billion
Portfolio Indicators			
Change to Total RDT&E costs from first estimate	37%	40%	42%
Change to Total acquisition cost from first estimate	19%	26%	25%
Total acquisition cost growth	\$183 Billion	\$301.3 Billion	\$296.4 Billion
Share of programs with 25% increase in program acquisition unit cost growth	41%	44%	42%
Average schedule delay in delivering initial capabilities	18 Months	21 Months	22 Months

Unfortunately, the scenario does not fare much better for space acquisition either, as space systems cost estimates have risen dramatically. For example, the US Government is already expecting a \$10.3 billion cost increase for the years 2009-2013 from original estimates (Chaplain, 2009). Obviously, these cost increases divert money away from new or existing space programs, making a cash-strapped environment for dollars even more competitive. Figure 1 displays the dramatic rise in cost and schedule delays for specific space acquisition programs.

Fiscal year 2009 dollars in millions



Source: GAO analysis of DOD data.

Figure 1. Differences in Total Life-Cycle Program Costs from Program Start and Most Recent Estimates
(GAO, 2009)

The GAO has identified specific issues causing significant cost increases and schedule delays as well as presented several best practices they believe might alleviate them. First, the GAO found many acquisition programs begin with an over dependence on emerging technology. If an unexpected delay occurs in the development of the technology, it negatively impacts the program schedule (Chaplain, 2009). Second, many programs attempt to fulfill all requirements in a single step—that is to suggest that traditional acquisition approaches will not work given the current user demands and timelines coupled with complex capability. “Programs have historically attempted to satisfy all requirements in a single step, regardless of the design challenge or the maturity of the technologies necessary to achieve the full capability” (Chaplain, 2009). Finally, the state of the Space Industrial Base is questionable. First-tier contractors are showing only marginal profit revenues and can absorb losses as they are diversified across the acquisition market. However, a recent Air Force Research Laboratory (AFRL) survey suggests that second- and third-tier contractors are producing miniscule profits. Consequently, many of these contractors are disappearing or leaving the space industry for more lucrative markets.

Dependence on Emerging Technology

The GAO found that the DoD has a tendency to fund more programs than they can afford. As a result, many programs come into existence already underfunded. In fact, in a survey presented to space program managers asking what are the top obstacles to achieving success, over 36% stated unstable funding (Chaplain, 2006). Additionally, programs that are performing well often find funding transferred to an underperforming program. This creates a very competitive environment for both program managers and contractors. In an attempt to secure as much funding as possible, a program manager will often over promise on capability and focuses on “bleeding edge” technologies that seem to demonstrate “the most bang for the buck” to those in control of the purse strings.

Unfortunately, this reliance upon bleeding edge technology comes at a significant cost. The GAO has found that developing technology in conjunction with program development is often fraught with schedule delays and cost overruns because the relied-upon technology often doesn’t work out as intended. In the same survey that measured unstable funding as a major obstacle, 18% of space program managers interviewed admitted that they relied on immature and untested technology (Chaplain, 2006).

Specific examples of space programs relying upon immature technology include Space Based Infra-Red (SIBRs) satellite system, the Advanced Extremely High Frequency (AEHF) satellite system and the National Polar-orbiting Operational Environmental Satellite System (NPOESS). In some cases, the GAO noted that sensors had not been fully tested or even prototyped, or software needs in general were just greatly underestimated (Chaplain, 2006). This reliance upon immature technology can, in some part, account for the fact that these programs have experienced a cost increase of 40% or more since their original cost estimation (Chaplain, 2006). In the case of SBIRs, their costs have more than doubled (Chaplain, 2006).

The GAO has provided two primary recommendations to help alleviate the overdependence upon immature technology. First, based upon best industry practices, the GAO recommends separating technology development from acquisition development. Implied in this concept is that program managers will have to rely upon and select proven technology. This leads to their second recommendation, which is to apply the technology readiness level (TRL) scale that the DoD “borrowed” from NASA. This rating scale evaluates technologies against 9 different levels of developmental maturity. The first level represents the least mature technology, whereas the ninth level represents the highest level of technological maturity. In the case of space acquisition, the GAO is recommending that programs only select and use technology that has reached a technology level of seven. For this particular level, the technology should be tested at least once in an operational environment, which is defined as “an environment that addresses all of the operational requirements and specifications required of the final systems” (USD(AT&L), 2002). Not surprisingly, the DoD has pushed back on this recommendation and offers its own recommendation of selecting technology that has reached level six. This level demonstrates a prototype or model of a system or subsystem in a relevant end-to-end environment, which is described as a test environment that simulates the key aspects of the operational environment (USD(AT&L), 2002).



Requirements and the Grand Design

There are several areas of concern with regard to space acquisition and requirements. First, as was alluded to earlier, many programs attempt to satisfy all users requirements in a single, grand design. Second, users and stakeholders have demonstrated a tendency to frequently change or add new requirements as the program is developing. Not surprisingly, these factors undoubtedly impact cost and schedule. According to Dr. Rustan, Director of Mission Support Directorate, NRO,

Our requirements-driven stakeholders often do not understand the cost implications of the various elements of their respective wish lists, and when we proceed to blindly integrate these capabilities, considerable problems develop. This problem is exacerbated when we are asked to hold fixed performance, cost and schedule at the beginning on any space acquisition, thereby inexorably increasing program risk. (Rustan, 2005)

Attempting to satisfy all stakeholders with a “grand design” is not only costly and has a long lead development time but also ineffective and perhaps impossible. Each stakeholder comes with their own needs and capabilities, which often come in conflict with other stakeholders requirements. For example, United States Southern Command (USSOUTHCOM) may come to the NRO and ask for an electrical optical satellite that can penetrate through dense jungle foliage. Whereas, United States Pacific Command (USPACOM) may ask for a satellite capable of performing broad area searches so they can identify ships in transit. Finally, US Central Command (USCENTCOM) may ask for satellite that can identify thermal signatures emanating from caves in an attempt to locate terrorist safe holds. Thus, designing this satellite becomes inherently complex because different technologies are required to achieve each of those capabilities. Not surprisingly, integrating these technologies so that they work together not only increases complexity but also increases risk to the budget and schedule.

Stakeholders also have a tendency to add or change requirements during the development of the system. According to Dr. Rustan, this may be attributed to new technology that has caught the user’s eye. Dr. Rustan also states that these stakeholders are often attempting to solve dynamic problems, problems that change and evolve rapidly over time (Rustan, 2005).

Spiral or Evolutionary development may well help alleviate some these problems. These acquisition strategies offer several advantages that seem to address the problems that are manifesting themselves in space systems development. More specifically, evolutionary acquisition acknowledges that not all requirements in a program are known up front. Therefore, evolutionary acquisition seeks to divide a program into more manageable increments or spirals in which the requirement can be more tightly defined. “Evolutionary acquisition’s primary goal is to reduce product cycle times by dividing and phasing requirements to produce initial capabilities sooner ... EA addresses requirements and technology risks by allowing requirements to evolve over time” (Ford & Dillard, 2009). Thus, these strategies allow program managers to take a complex design and break it down into smaller, more manageable intervals. Additionally, they allow the program manager to provide an initial set of capabilities to the user sooner than traditional acquisition approaches. The ability to deliver capabilities may also help negate the effects of requirements creep as well.

However, there is new evidence suggesting that evolutionary development may actually be costlier and longer. More specifically, modeling and simulation performed by



David Ford and John Dillard suggests that although spiral development may satisfy first phase requirements faster, it certainly is more expensive when compared to traditional acquisition approaches. Further, their modeling also suggests that it takes longer to satisfy all requirements when compared with traditional approaches (Ford & Dillard, 2009). See Table 2 for their results. Therefore, it would appear a comprehensive evaluation is needed to determine if the best practices recommended by the GAO will truly fix space acquisition.

Table 2. Performance Comparison of Three Simulated Acquisition Projects
(Ford & Dillard, 2009)

Performance Measure	Units of Measure	Project Scenario		
		<i>Javelin</i>	<i>Base Case Traditional</i>	<i>Base Case Spiral</i>
Duration to first requirement satisfied	Weeks	471	470	397
Duration to maximum requirements satisfied	Weeks	520	518	762
Total development cost	\$1.0 Million	722	719	1,555
Requirements satisfied by deadline	Percent	100	91	18
Final requirements satisfied	Percent	100	91	91

State of Space Industrial Base

A robust science, technology, and industrial base is critical for US space capabilities ... departments and agencies shall ... encourage an innovative commercial space sector, including the use of prize competitions; and ensure the availability of space related industrial capabilities. (Office of Science, 2006)

When examining the health of the space industrial base there seems to be two primary issues. First, is the *International Traffic in Arms Regulation (ITAR)* export regulations controlled by the US State Department? Second, is their large dependence upon the US Government for business and revenue? Specifically, 60% to 65% of sales for the space industrial base were from the US Government between the years 2003 to 2006 (Chao, 2008). In either case, while the first-tier contractors (Northrop Grumman, Lockheed Martin and Boeing) are showing minimal profit margins, the second- and third-tier contractors are leaving the industry. They are either going out of business or don't feel there is enough of a business case to continue to engage in space acquisition.

ITAR is designed to prevent protected technology and munitions from transferring to other nations and to enable the US to maintain its technological superiority. It is administered by the US State Department and was brought about as a result of Chinese technological gains from their observation of US investigative techniques on the failed ASTAR II launch (DoD, 1998). Despite its good intentions, strong arguments are arising that it is failing to prevent space technology from developing in other nations and is hurting the US's technological superiority.

For example, many foreign competitors, such as Thales, are advertising their satellites as *ITAR* free. The benefit provided here is that foreign companies don't have to



progress through a complex and confusing export license process to acquire similarly capable components. Additionally, many nations when issuing a request for proposal purposely restrict the response to their proposal to less than sixty days. This virtually eliminates US firms without having to worry about economic retaliation because US firms first have to apply for an export license under *ITAR* before they can compete. According to a recent survey conducted by AFRL, the average turn around time for a license approval was 106 days in 2006 (Chao, 2008). Thus, it would appear that *ITAR* is further encouraging US firms to remain dependent upon the government as it is increasingly difficult to enter and compete in international markets.

It would also seem that *ITAR* has encouraged foreign nations to develop their own space technology. For example, we are seeing greater cooperation among foreign competitors in space, particularly among the Europeans. Further, China is closing the space gap with the US. They have developed their own positional navigational system, conducted their first manned space flight, demonstrated a successful space walk and successfully tested anti-satellite weapons technology (Chao, 2008). This would suggest that *ITAR* is not achieving its objective of maintaining US technical superiority in space.

The Department of Defense (DoD) provided a report to the Congress in February 2006, stating that the aerospace and defense industry was outperforming companies listed on the S&P 500 (DeFrank, 2006). As a whole, the industry did seem to outperform the S&P 500. However, when looking specifically at the space industry, and especially the second- and third-tier suppliers, we see that these companies had significantly lower profit margins. A recent survey conducted by Air Force Research Laboratory shows that second- and third-tier suppliers were only bringing in profit margins of 4% to 6% (Chao, 2008). These low profit margins mean there is less revenue to invest in their personnel and in their research and development. Combined with pressures from prime contractors to provide the “best possible price,” these companies become less competitive, and, thus, we see a “hollowing out [of] the supply chain” (Chao, 2006). Additionally, the Suppliers Excellence Alliance asserts that 50% of all second- and third-tier suppliers will cease to exist by 2011 (Chao, 2006). This represents a serious problem because the primary contractors subcontract out approximately 80% of their space acquisitions to these lower-tiered companies (Chao, 2008).

The Center for Strategic International Studies (CSIS) provides several recommendations that may help alleviate the current situation. First, the Department of State should conduct a technological review to determine which technologies are critical and which ones are not (Chao, 2008). However, simply labeling a satellite as a critical technology in its entirety seems counterproductive. Currently, the US is the only nation in the world that labels a satellite as a munition, whereas many European nations designate them as dual use technology. Second, provide authority to those entities involved in satellite exports to “review cases in a real time, case by case, specific time period” (Chao, 2008). Finally, those government entities involved in space acquisition should annually review the state of the space industrial base.

Summary

Although space acquisition originally provided the US with an advantage in the Cold War, it has also impacted nearly every aspect of our society. It plays a role in our banking and financial industries, provides entertainment, enables us to communicate half-way around the world, provides information on otherwise denied areas, and acts as a force multiplier in the conduct of military operations. The Persian Gulf War demonstrated the



asymmetrical advantage that space systems provided for US forces. As a result, other nations made a series of steps and commitments to develop their own space assets. As the US became more reliant upon satellite technology, its users began expressing requirements for a greater number of systems with far greater capabilities.

System complexity is only one aspect affecting US space acquisition today. There is a large body of literature that suggests US space acquisition is over budget, behind schedule, and delivering underperforming systems. The GAO seems to attribute a number of factors to contributing to this situation. However, three primary factors include an over-reliance on immature technology, managing requirements to build the “grand design” and the health of the space industrial base.

Addressing these factors will be critical so that the US can maintain its technology superiority and leadership in space. This is especially critical as countries such as Russia and China continue to mount significant challenges to our dominance in space. A loss of US leadership in space could very well translate into a loss of prosperity and national security.

References

- Azani, C., & Flowers, K. (2005, January-February). Integrating business and engineering strategy through modular open systems approach. *Defense AT&L*, 37-40.
- Blakey, M.C. (2009, April 30). Testimony to Committee on House Armed Services Subcommittee on Strategic Forces, Space Systems Acquisition Statement.
- Chao, P. (2008, February 19). *Health of the US space industrial base and the impact of export controls*. Center for Strategic International Studies.
- Diamond, J.M. (2001). Re-examining problems and prospects in US imagery intelligence. *International Journal of Intelligence and Counter-Intelligence*, 14(1), 1-24.
- Dillard, J.T. (2004, August-November). Toward centralized control of defense acquisition programs. *Defense Acquisition Review Journal*, 331-344.
- Eiband, D. (2005, August-November). Innovative procurement strategies. *Defense Acquisition Review Journal*, 323-329.
- Ford, D.N., & Dillard, J. (2009, July). Modeling the performance and risks of evolutionary acquisition. *Defense Acquisition Review Journal*, 143-158.
- Fowler, D.N. (2001, Fall). Innovating the federal acquisition process through intelligent agents. *Acquisition Review Quarterly*, 151-166.
- GAO. (2003, September). *Improvements needed in space systems acquisition management policy* (GAO-03-1073). Washington, DC: Author.
- GAO. (2004, January 29). *Defense acquisitions risk posed by DOD's new space systems acquisition policy* (GAO-04-379R). Washington, DC: Author.
- GAO. (2005, June 23). *Defense acquisition incentives and pressures that drive problems affecting satellite and related acquisitions* (GAO-05-570R). Washington, DC: Author.
- GAO. (2005, July 12). *Space acquisitions: Stronger development practices and investment planning needed to address continuing problems* (GAO-05-891T). Washington, DC: Author.
- GAO. (2006, April 6). *Space acquisitions: Improvements needed in space systems acquisitions and keys to achieving them* (GAO-06-626T). Washington, DC: Author.
- GAO. (2006, June 1). *Defense acquisitions: Space system acquisition risks and keys to addressing them* (GAO-06-776R). Washington, DC: Author.
- GAO. (2006, September). *Space acquisitions: DOD needs to take more action to address unrealistic initial cost estimates of space systems* (GAO-07-96). Washington, DC: Author.



- GAO. (2009, April 30). *Defense acquisitions: Charting a course for lasting reform* (GAO-09-663T). Washington, DC: Author.
- GAO. (2009, May 20). *Space acquisitions: DOD faces substantial challenges in developing new space systems* (GAO-09-705T). Washington, DC: Author.
- GAO. (2009, May 30). *Space acquisitions: Government and industry partners face substantial challenges in developing new DOD space systems* (GAO-09-648T). Washington, DC: Author.
- Gholz, E. (2007, February). A business model for defense acquisition under the modular open systems approach. *Defense Acquisition Review Journal*, 217-233.
- Hamel, M.A. (2009, March 10). *America's leadership in space*. Memorandum for the President, Committee for US Space Leadership.
- Hartman, J. (2009, April 30). Testimony to Committee on House Armed Services Subcommittee on Strategic Forces, Space Systems Acquisition Statement.
- Jakovelić, M. (2006). Modular open system approach (MOSA) and TTP-based platforms for aerospace control systems. *IEEE*.
- Kohler, R.J. (n.d.). *The decline of the National Reconnaissance Office*. Center for the Study of Intelligence. Retrieved June 3, 2008, from <https://www.cia.gov/library/center-for-the-study-of-intelligence/kent-csi/docs/v46i02p.htm#author>
- Levin, C., & McCain, J. (2009, April). *Weapon system acquisition reform act 2009*. Washington, DC: US GPO.
- Mazur, J., Jr. (2003, Winter). Acquisition secrets from the National Reconnaissance Office. *Air Force Journal of Logistics*, 27(4), 4-13.
- NRO Chief Operating Officer (COO). (2009, April 3). *Integration, verification and transition procedure* (CIVVTP).
- Office of Science and Technology, Executive Office of the President. (2006, August 31). *US National Space Policy*.
- Pagliano, G.J., & O'Rourke, R. (2004, April 8). *Evolutionary acquisition and spiral development in DOD programs: Policy issues for Congress*. CRS Report for Congress.
- Program manager's guide: A modular open systems approach (MOSA) to acquisition*. (2004, September). Open Systems Joint Task Force. Retrieved from <http://www.acq.osd.mil.ostfj>
- Rendon, R.G. (2007). Using modular open systems approach in defense acquisitions: Implications for the contracting process. *IEEE*.
- Richelson, J.T. (2002). Restructuring the NRO: From the Cold War's end to the 21st century. *International Journal of Intelligence and Counter-Intelligence*, 14, 496-539.
- Rustan, P.L. (2005, July 12). Testimony to the United States House of Representatives, House Armed Services Committee Hearing, Space Acquisition.
- Shannon, M.J. (1994, June). The Clementine Satellite. *E&TR*, 1-11.
- Spring, B. (2005, October 19). *Congressional restraint is key to successful defense acquisition reform* (No. 1885). The Heritage Foundation.
- USD (AT&L). (2002, April). *Mandatory procedures for major defense acquisition programs (MDAPS) and major automated information system (MAIS) acquisition programs* (DoD 5000.2-R). Washington, DC: Author.



Panel #22 – System Engineering for Project Success

Thursday, May 13, 2010

1:45 p.m. –
3:15 p.m.

Chair: James Wolfe, Head, Strategic and Weapon Control Systems Dept,
Naval Surface Warfare Center Dahlgren

***Systems Engineering Applied Leading Indicators—Enabling Assessment
of Acquisition Technical Performance***

Paul Montgomery and Ron Carlson, Naval Postgraduate School

Funding for Life: When to Spend the Acquisition Pot

Kate Gill and Kirsty Carter-Brown, Defence Science and Technology
Laboratory

***System Capability Satisficing in Defense Acquisition via Component
Importance Measures***

Brian Sauser, Jose Ramirez-Marquez, David Nowicki, Weiping Tan,
Romulo Magnaye, Stevens Institute of Technology; and Abhi
Deshmukh, Texas A&M University



Systems Engineering Applied Leading Indicators— Enabling Assessment of Acquisition Technical Performance

Paul Montgomery—After retiring in 1990 from a twenty-year career in the Navy, Dr. Montgomery served as a Sr. Systems Engineer with Raytheon and Northrop Grumman corporations and developed communications, surveillance, and sensor systems for commercial, military (USN, USA, USAF), and intelligence communities (NSA, NRO). He earned his doctorate in Systems Engineering from George Washington University (DSc '07), performing research related to cognitive/adaptive sensors, MSEE (1987) from the Naval Postgraduate School, and BSEE (1978) from Auburn University.

Dr. Paul Montgomery, NPS
Associate Professor of Systems Engineering
prmontgo@nps.edu
703 568 1165

Ron Carlson—Carlson served twenty-six years in Naval Aviation as a pilot, seven years of which were at NAVAIR, where he led NAVAIR Systems Engineers through several years of systems engineering revitalization to the NPS SE department. He is currently in the Systems Engineering doctoral program at Stevens Institute of Technology, earned Master's degrees in Strategic Studies and National Policy from the Naval War College and Business Administration-Aviation from Embry Riddle Aeronautical University, and his Bachelor of Science in Nuclear Engineering from the University of Michigan.

Ron Carlson, NPS
Professor of Practice of Systems Engineering
rrcarlson@nps.edu
301 757 9743

Both Dr. Montgomery and Ron Carlson are SE department embedded faculty members, providing onsite research and instruction support to NAVAIR (Patuxent River, MD), NAVSEA (Dahlgren, VA, Carderock, MD), and other east coast NPS SE students.

Abstract

This paper discusses research in developing DoD acquisition metrics associated with Systems Engineering activities that may provide greater insight into the technical performance of development programs. These metrics are called Systems Engineering Applied Leading Indicators (ALI). We examine current development of single and multi-factor ALIs that have been developed during the past year at the Naval Air Systems Command (NAVAIR) in Patuxent River, MD. The development methods, early examination of ALI utility, and user acceptance are discussed. The authors have been embedded with the NAVAIR Systems Engineering Development and Implementation Center (SEDIC) (the center of this work for NAVAIR) as part of this ALI exploration.



Acknowledgments

The authors have been working in close collaboration with the members of the NAVAIR SEDIC team at NAS, Patuxent River, MD, in the conduct of the subject research. Much of the foundational work referred to in this paper is a result of their efforts. The team includes Juan Ortiz, Paul Hood, and Javier Sierra of NAVAIR, and Gregory Hein, Johnathan Gilliard, and Tina Denq from Booz Allen Hamilton.

Introduction and Problem Definition

Background

What is the role of systems engineering (SE) in the acquisition and development of systems? The professional society for SE (INCOSE) defines SE as follows:

*Systems engineering is an interdisciplinary approach and means to enable the realization of successful systems. It focuses on defining customer needs and required functionality early in the development cycle, documenting requirements, and then proceeding with design synthesis and system validation while considering the complete problem: operations, cost and schedule, performance, training and support, test, manufacturing, and disposal. **SE considers both the business and the technical needs of all customers with the goal of providing a quality product that meets the user needs.*** (INCOSE, 2010)

The principles, practices, and methods of SE are well defined and long practiced by Government and industry (INCOSE, 2010; NASA, 2007; NAVY, 2004). The value added by disciplining the development of a system is well appreciated. In the mid 1990s, SE practices were augmented with the concepts of SE metrics (INCOSE, 1995; 1998; Roedler, 2005). Early implementation of these metrics has been directed at the measurement of the performance of the SE process itself.

In the bolded part of the definition above, SE continues to expand its benefits to include not only the development team but also all customers and stakeholders who are maximally interested in a project/program that is delivered satisfying cost, schedule, as well as technical goals. There is now interest in the SE community on how to expand, define, and derive metrics and methods that would provide predictive or prognostic indicators of the success of a development effort as a whole (see **Error! Reference source not found.**) (Rhodes, Valerdi & Roedler, 2009). While the existing SE metrics and methods have typically produced lagging and inferred indicators of the health and status of a development effort, current efforts and research are now underway to examine how to provide direct *leading indicators*, derived from SE and *applied* to understanding and predicting the technical trajectory of the aggregate development effort. Because we are *applying* and focusing the concepts of SE leading indicators (Roedler & Rhodes, 2007), we will refer to this concept as SE Applied Leading Indicators (ALI) for the remainder of this paper.



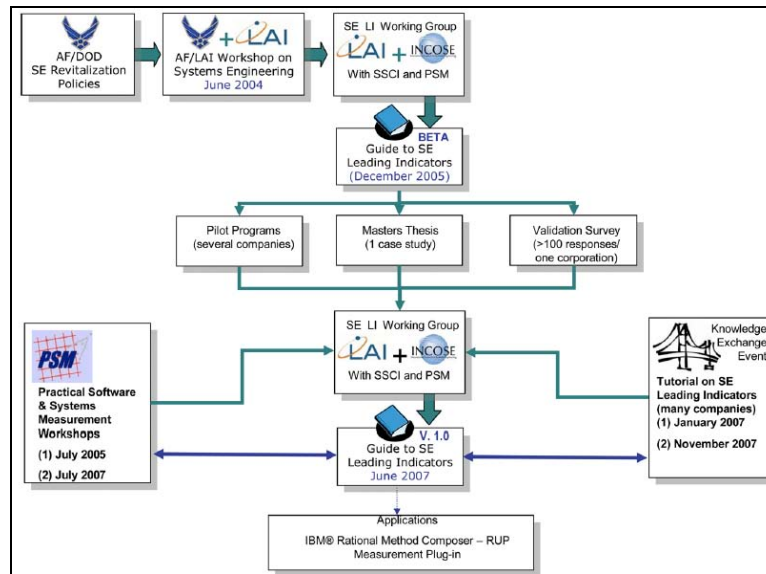


Figure 1. Government/Industry Partnership Exploring SE Leading Indicator Concepts and Application
(Rhodes et al., 2009)

The authors set out attempting to focus on why programs fail to meet user expectations at delivery. Our goals were to determine what engineering metrics could be defined and analyzed to provide such insight where programs are apparently not getting such insight today (based upon failure rates of system qualification testing results). This goal lead us to intersect ongoing efforts related to SE ALIs that we determined would provide an understanding of closely related metrics and processes that we would underpin our investigation. The authors have been supporting and co-researching with Naval Air Systems Command (NAVAIR) in Patuxent River, MD, to examine the identification, relevance, and application of SE ALIs. NAVAIR has been examining the ALI concept through engagement with acquisition offices, data gathering and analysis, formulation of predictor algorithms, and prototype ALI tool development. The Systems Engineering Development and Implementation Center (SEDIC) is conducting this NAVAIR effort in collaboration with working groups depicted in Figure 1.

Problem Definition

Program managers apply well-proven and refined program metrics and control mechanisms largely based upon Earned Value Methods (EVM). The EVM fulcrum metrics are cost and schedule that generate analysis outputs depicting variances from plans and estimates. From EVM analysis, program cost and schedule status can be assessed and projection of those parameters can be inferred. Program managers, however, are not provided abundant metrics that can provide insights into the technical health of a development effort and indications of the trajectory of program health, good or bad. Risk metrics and processes provide some indications of technical health but are often qualitative and provide little algorithmic opportunities for prognostics. In general, program managers are faced with the development of complex systems, use EVM and risk management effectively; however, programs are failing to fully control costs and can routinely exceed cost estimates by 25% or more (see Figure 2).

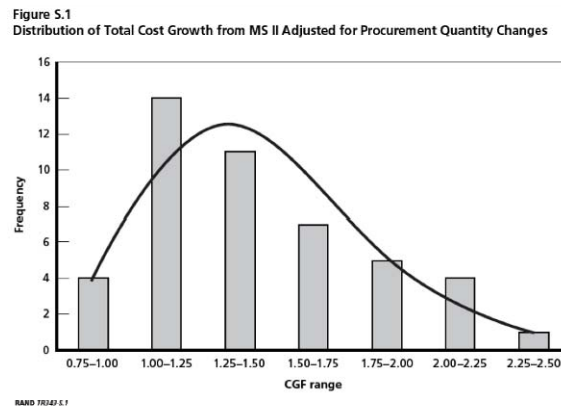


Figure 2. Control of Cost Growth of Programs Remains a Challenge
(Arena, Robert, Murray & Younossi, 2006)

In addition to the quantity of programs that exceed cost estimates, it appears that acquisition cost growth can be attributed to causes centered upon control of technical baselines (see Figure 3). The development of ALIs is intended to gain much more granular insight into the development of the technical baselines as soon as possible to allow for both assessment and predicted program performance so mitigation can be applied. In summary, the specific problem and research response follows:

Problem—Program managers do not have access to adequate technical metrics in order to provide high fidelity assessment of technical health of a complex system development program and quantitative prediction of technical performance.

Research Question—Can SE technical metrics be identified, quantified, and methodically applied to complex system developments to provide technical assessment and leading indications of technical program performance and ultimate success?

Research Objectives:

- Identify relevant data supporting the development of ALIs,
- Identify leading indicators tailored to systems engineering effectiveness,
- Prototype ALI user tools to measure relevance, acceptance, and obtain feedback, and
- Identify new, revised, or derived metrics to support refined ALI methods.

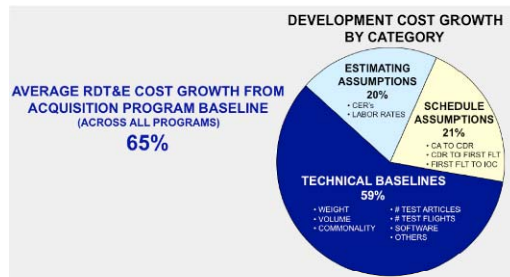


Figure 3. Cost Growth Largely Impacted by Control of Key Attributes of Technical Baselines
(NAVAIR 4.2 Division Analysis)

Advance Leading Indicator Concepts

Technical Measurements

SE processes provide metrics, measurements, and analysis activities throughout systems development. These technical measurement activities provide lead system engineers and project managers insight into project technical performance and associated risks. These metrics are most often associated with Measures of Effectiveness (MOEs), Measures of Performance (MOPs), and Technical Performance Measures (TPMs) with associated Key Performance Parameters (KPPs) and Key System Attributes (KSAs). These associations and metrics are qualified through continual testing and often manifest themselves graphically using control chart methods (see Figure 4).

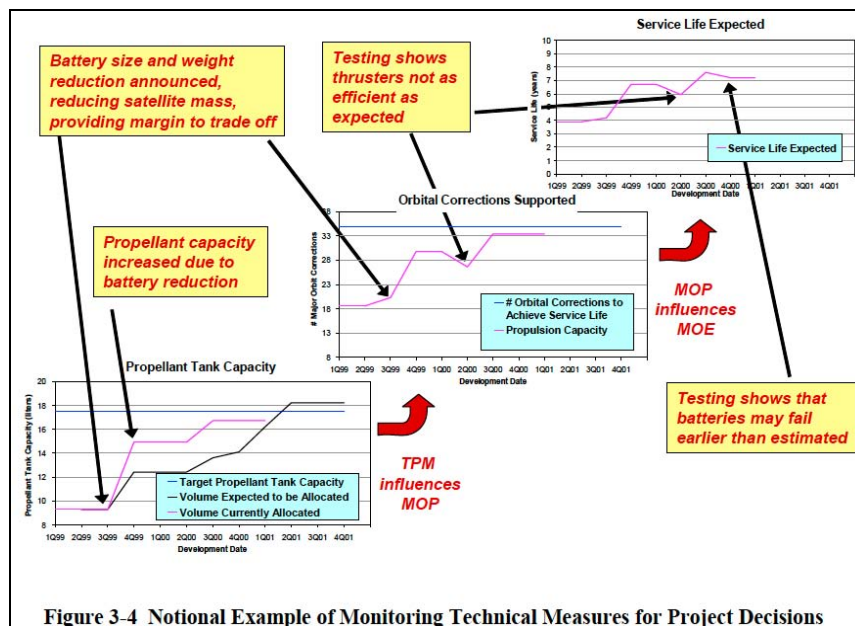


Figure 4. Technical Measures Associated with MOEs, MOPs, and TPMs Guide Testing and Achieving Specifications
(Roedler, 2005)

The above technical measurement processes are often focused on assessing the progress of the system in meeting specification as development unfolds. Although the development of ALIs seems similar to these practices, the intent of ALIs is provide a more holistic and prognostic assessment of the technical aspects of the project by integrating both system technical metrics as well as systems engineering-derived process metrics. ALIs, although substantiated in historical performance of similar projects, are highly forward-looking and technical-rich in fidelity. They are intended to inform the project technical approach and be fully integrated with the program management approach (see Figure 5).

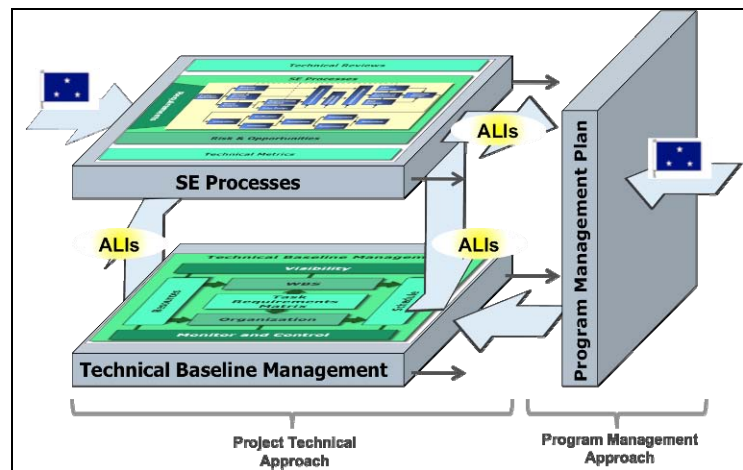


Figure 5. ALIs Provide Metrics Rooted in SE Technical Approach and Supports Program Management Approach

The development and use of ALIs are intended to augment existing program/project management methods, not replace them. Although influenced by many similar metrics (e.g., cost, schedule, etc.), ALIs are derived from system attribute and system engineering metrics to produce technical health and prognostics that enhance the program manager's overall assessment and direction of the project (see Figure 6). They enrich the existing EVM-derived assessment to provide project leadership higher fidelity project technical status and direction that enables greater decision analysis completeness.

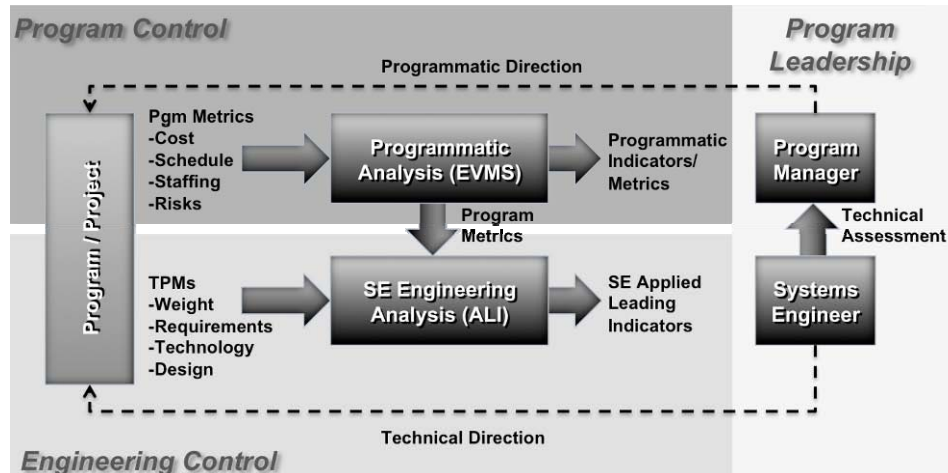


Figure 6. SE Applied Leading Indicators (ALI) Augment Program Management

ALI Approach

Data Selection and Collection

The development of ALIs has progressed through multiple phases. The initial phases identified metrics that were not being integrated into technical assessment that could be derived from engineering activities. The ALI effort is careful to augment program management evaluation methods, not duplicate them. The identification of metrics led to the following list of relevant technical metrics that were considered readily available in program data repositories:

- Aircraft empty weight,
- Software metrics,
- Architecture metrics,
- Requirements metrics,
- Technical review closure burn down rates,
- Reliability, Availability, and Maintainability (RAM) metrics,
- Technical risk metrics,
- Engineering staffing metrics,
- System complexity, and
- Technology maturity.

The data above is collected to form a historical baseline of program performance of similar or related programs (later ALI phases would incorporate current program data to predict future performance). The focus of early research has been associated with the technical metric of *aircraft weight* and subsequent modeling impacts of aircraft weight-growth to program cost-growth throughout development. This data was collected from historical records from the program division records. The data was also “affinitized” or grouped in like-program categories to maintain relevance of analysis results. Examples of these groupings included aircraft development with similar plan forms (e.g., rotary, fixed

wing, remotely piloted, etc.), size of program (ACAT I, II, etc.), and mission (fighter, transports, etc.). In all, approximately 11 programs form the foundation for further data analysis.

ALI Data Analysis

Armed with a single dimension of ALI applicability (weight-growth versus cost-growth), analysis has been conducted to assess what statistical methods can be applied to this data with confidence and with assurance that the methods are representative, relevant, and extensible to provide useful data to program managers. The steps for the early “single-factor” ALI analysis is depicted in Figure 7 and is summarized below.

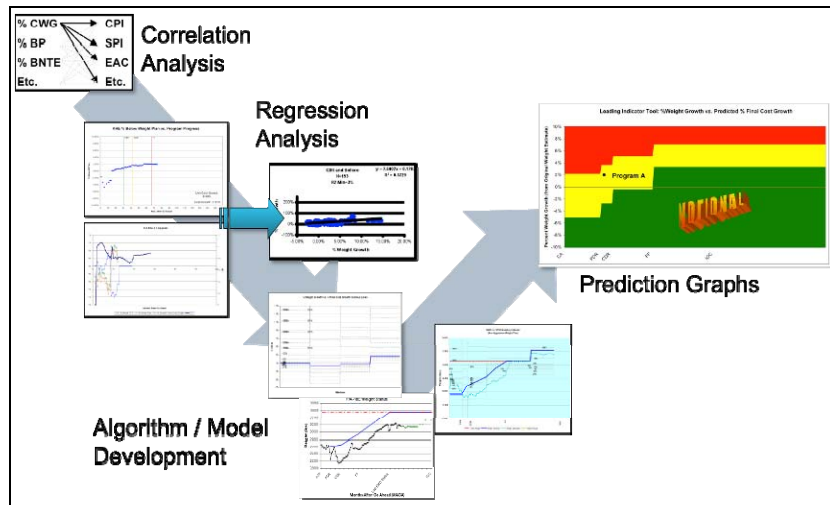


Figure 7. ALI Data Analysis Transforms SE Historical Metrics to Algorithmic, Leading Assessment of Technical Performance and Prototype ALI Tool Development

- **Test for impact (correlation)**—Of several program performance impacts related to aircraft weight throughout development, cost growth proved to be highly correlated and validated and formed the basis for continued analysis.
- **Test for meaningful relationships (regression)**—The ability of weight growth to predict cost growth was examined through several regression methods. The results showed significant statistical strength of using weight-growth as a cost-growth predictor; however, the data must be segmented into major epochs of program development to maximize this predictive strength. The epochs were separated by major design reviews (e.g., PDR, CDR, First Fight, etc.) to ensure predictive usefulness.
- **Derive algorithms to develop prognostics**—The segmented data produced regression formulations with confidence and fit parameters for each epoch and program types. These formulations established historical performance baselines.
- **Validate models against all programs**—The data was validated against the affinitized groupings to ensure that the regression models derived above provided an accurate model of historical performance of the programs they modeled.
- **Develop “tripwire” boundaries prediction zones**—Although a prediction algorithm was considered useful, it was considered most valuable if cost-growth boundaries could be added to the prediction. These boundaries could be

adjusted based upon the interest of the program manager, but as a minimum, would be set at cost-growth conditions that would alert the program manager and leadership of severe program trouble (e.g., 10 USC 2433 (Nunn-McCurdy) limits). These limits were segmented into classic color zones (or “tripwires”) of assessment (e.g., red, yellow, and green).

- **Build tool for user validation**—The predictive regression-based formulation was combined with the tripwire assessments/graphics to prototype a user tool for inputting and assessing their program based upon current weight estimates and comparing their current weight estimate (at a particular time of the development) and comparing to historical performance of similar programs to provide a predictor of program cost-growth performance

Early Single-factor ALI Feedback

The user community (program managers of NAVAIR acquisitions) were presented the output of the tool is depicted in **Error! Reference source not found.** to obtain feedback related to ALI concept and utility. The graphic is described in the following paragraph.

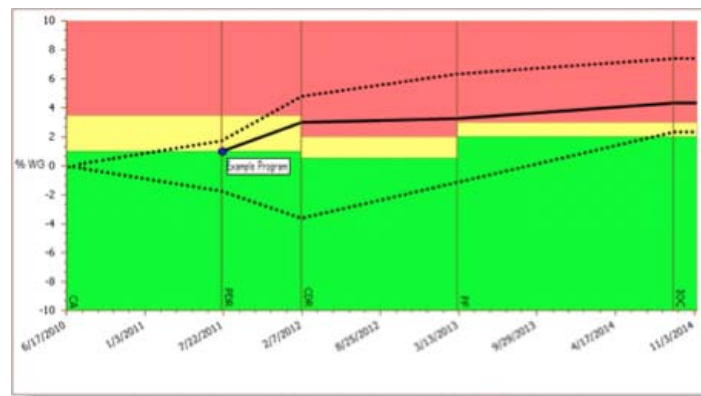


Figure 8. ALI Prototype Tool Provides Cost-growth Prediction Based Upon Similar Program Histories as well as Status of Subject Program
(Denq, Hein & Gilliard, 2009)

The starting “dot” in the yellow zone is the current assessment of an example program. The solid black line is the predicted cost growth of the program based upon where the current program is in comparison to historical performance of similar programs. The dotted lines are confidence bounds of the prediction. This graph, once again, is the depiction of the impact of a single ALI (weight growth) on program cost growth. Early program manager and lead engineer feedback revealed that this graph, although informative to a degree, generates more questions than it answers. Some examples include:

- If single-factor ALI analysis predicts cost growth, what other factors may also impact cost growth?
- What are the impact comparisons among single ALIs?
- Do other ALIs “mutual couple” to cause cost growth?
- What do I (PM/SE) do about it?
- How much is my program like historical programs?
- How can I input my own predictive performance judgment into the algorithm?

Moving to Multi-factor ALIs

The most generalized feedback from program managers and systems engineers to the early single-factor ALI concept is that it needs (1) to consider more ALIs, (2) incorporate their interactions, and (3) algorithmically combine their influences into an integrated ALI metric for the program. Similar to EVM integration of cost, schedule, and achievements (milestone completion) into a few key metrics, ALI needs to work toward that goal. The process for moving to an integrated ALI output is shown in Figure 9.

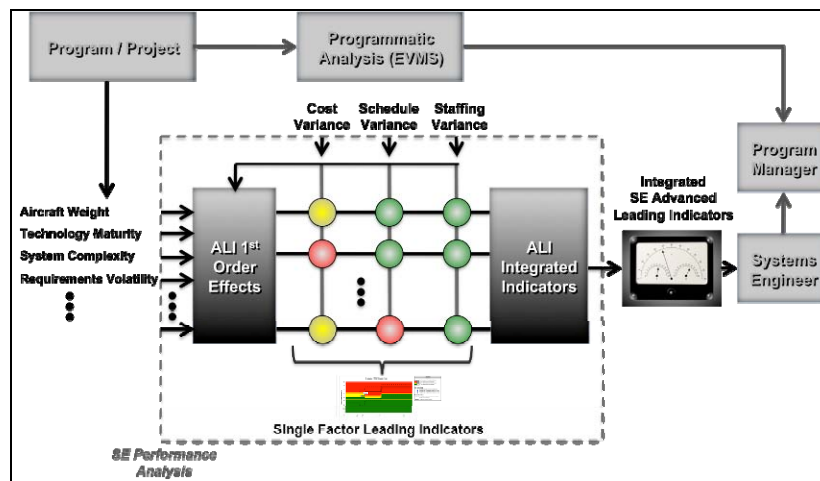


Figure 9. Single-factor ALI Analyses Are First Steps to an Integrated ALI Output

The single-factor ALI analysis and formulations are shown in the center of the diagram. They are analyzed individually and then, after model validation, are integrated to provide a more “global” ALI metric. The repeated analysis steps are depicted in Figure 10. This process has led to an attempt at an integrated, multi-factor ALI approach that is currently being explored.

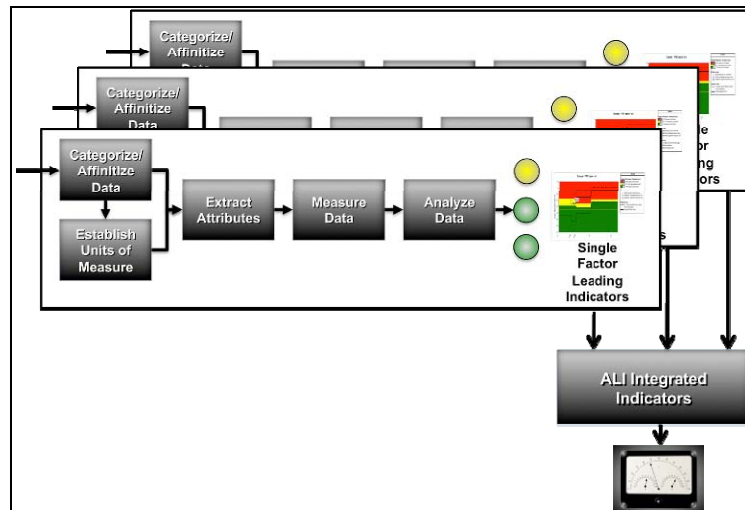


Figure 10. Parallel and Independent Single-factor ALIs Lead to an Integrated ALI for the Program

Multi-factor ALI Development

As discussed above, single-factor ALI development and research has led to the current research into multi-factor ALIs. The underlying assumption is that if a single-factor ALI concept was validated historically, proved some utility in prediction program performance, and had statistical salencies that could be exploited in a tool, then we may be able to ingest multiple ALI metrics simultaneously and provide meaningful analysis using related statistical methods suited for multi-factor analysis. Ongoing multi-factor ALI investigation (see Figure 11):

- Retains historical data analysis of key program ALI metrics (this maintains a credible baseline of program performance upon which to compare programs).
- Applies multiple regression methods.
- Integrates user assessment of both current conditions and their predictions of individual ALI future performance (e.g., if your program is currently 5% over weight, what is your prediction of how this metric will change in the future?).
- Applies program end-state simulations based upon historical formulation and user estimates. After establishing both historical baseline and associated multiple regression algorithmic models, user predictions are integrated into the models via simulations to predict program performance, fit, and confidence limits.
- Provides integrated multi-factor ALI graphical output to the program leadership.

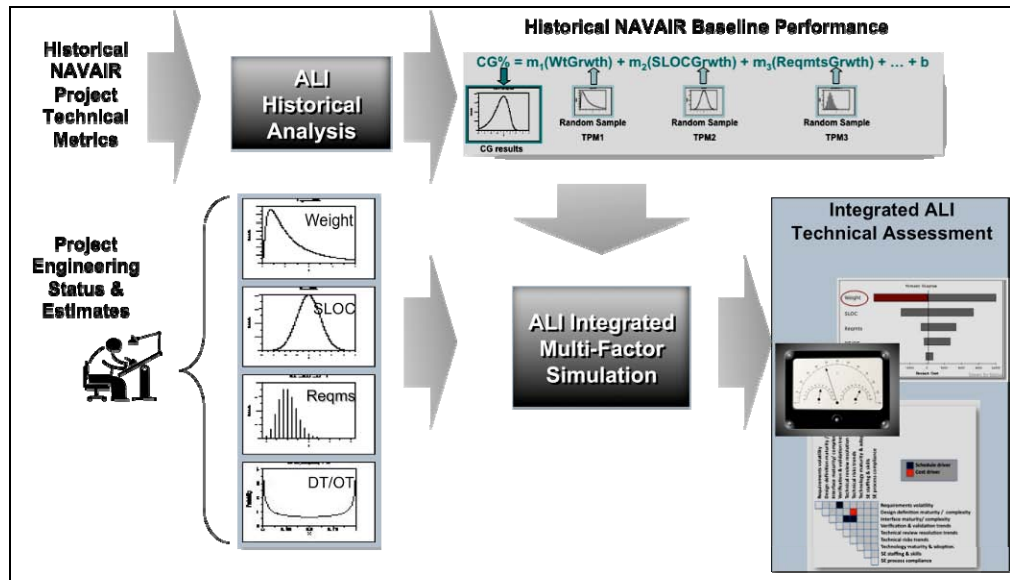


Figure 11. Multi-factor ALI Development/Research Approach

Early graphical concepts are intended to give insights into the “mutual coupling” among the ALIs and their impact on the program. Some concepts include an “interaction matrix” approach (Figure 12, left) showing, for example, which multiple ALIs drive program cost and schedule (indicated by colors) and provide insight into their possible interactions (inferred by their relationships vertically and horizontally). Additionally, from multi-factor ALI analysis, it may be possible to depict which factors are most influential on program performance

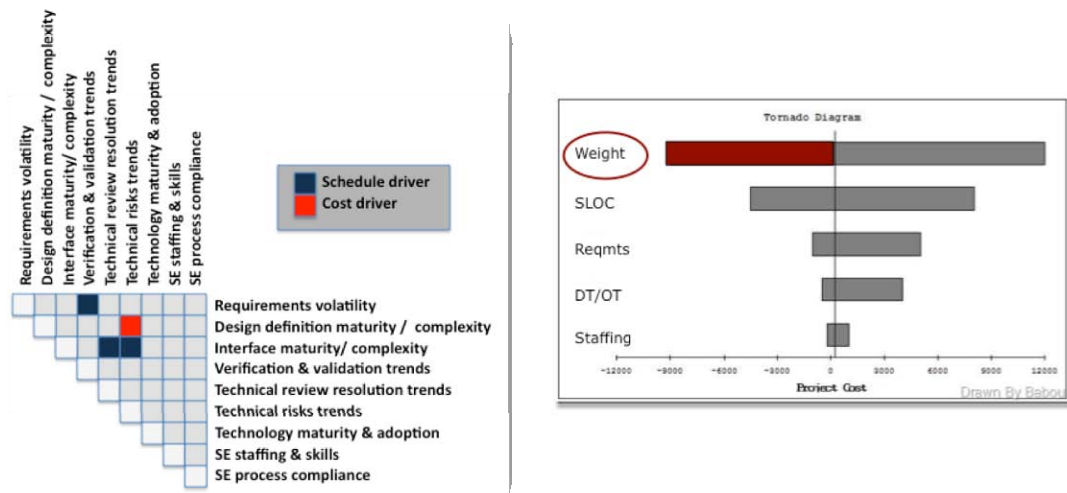


Figure 12. Example of Multiple Alis Influencing Program (Left) Cost and Schedule (Color) and Inferring Their Possible Interactions (Vertical/Horizontal Association) and (Right) Key ALI Influencers

ALI Insight into System Qualification Testing Success

Consistent with the authors' original goals, an NPS capstone project thesis investigated using the available ALI analysis data to gain insight into how programs were succeeding in their qualification testing (Buchanan & Jungbluth, 2010). Their research indicates some promising, although weak, statistical inferences about the data and successful testing outcomes. Their work sets foundations for further research discussed below.

Results and Conclusions

Although this ALI research is in the early stages, the ALI strategy, methods, and results discussed in this paper show promise for providing program manager and lead system engineer insight into the current and predicted technical success of their programs. This has been demonstrated through ALI data analyses, ALI user tool prototypes, and user acceptance testing.

This research began with a focus on why programs fail to meet user expectations at delivery. The goal is to determine what engineering metrics can be defined and analyzed to provide insight into success of qualification testing (e.g., operational test and evaluation, validation, etc.). This goal lead us to intersect ongoing efforts related to SE ALIs that we determined would provide an understanding of closely related metrics and processes that would underpin our investigation. The ALI research is still formative and evolving and the following conclusions are mostly qualitative (non parametric) but help to refine further directions related to ALIs and original research goals.

Data—Although there are rich data repositories available in the case of NAVAIR, the data can be inconsistent and incongruent. This increases difficulty in data analysis and bounding uncertainty in the predictive credibility of the ALI algorithms and tools. Additionally, retention of data from various programs is sometimes incomplete, leading to statistical analysis of sparse data. These problems are not, however, insurmountable and occur regularly in statistical analysis activities. The benefit of the ALI investigation is that recommended ALI metrics will emerge that can be recommended to be inculcated into the acquisitions to enable greater future ALI fidelity, granularity, and reliability.

Single-factor ALI analysis—The weight-growth versus cost-growth ALI analysis revealed that the development method was valid, provided a basis for ALI tool prototyping, and garnered preliminary user acceptance, understanding, suggested improvements, and identified ALI concept shortfalls. The technical basis is strong; however, the most impactful recommendation from users was to demand *multi-factor* ALI methods.

Note: When we tried a “programmatic” metric (staffing-growth versus cost-growth) as a comparison, the statistical predictive strength was not as strong as the technical metric of weight. The resulting conclusion was that there are many external factors (re-baselining, inter-program staff balancing, etc.), which weakened statistical fit. Additionally, although we have some interest in multi-ALI interactions with programmatic metrics, we discontinued the staffing investigation because it proved too parallel with programmatic metrics (i.e., EVM).

Multi-factor analysis—These methods and analyses are in very early stages. Early models and processes are employing data from the same programs, leveraging single-factor analysis lessons learned, expanding to include multivariate statistical methods, and new graphical output techniques. Early indications using simulated modeling data show promise.



The next steps will include actual data, validate multivariate models, and prototype a tool to garner user acceptance.

ALI metric expansion—The only metric that was validated was aircraft weight and its growth throughout the development cycle. More metrics still need to be developed and incorporated into the research.

User acceptance—Users recognize the need for a method based upon technical metrics to provide predictive program performance insight. They do not, however, want ALI to replicate EVM-based metrics and methods. Additionally, they desire ALI methods to incorporate prediction inferences and judgments of the project engineering and management team to influence analytical output. Finally, as stated earlier, user inputs showed a strong need to reveal mutual coupling of the multiple ALI factors, the overall impact to the program, and insights into how to respond, technically.

Areas for Continuing Research

Multi-factor ALIs—As stated above, this analysis is in the early phases and needs to be completed to the point of testing, validation, and user acceptance/feedback. The next steps are to include ingesting actual data, validating multivariate models, and prototyping a tool/user interface to gain insight into user acceptance

Total-ownership-cost control—During the conduct of this research, there is a “sea change” underway toward Total Ownership Cost (TOC) control at NAVAIR. This potentially shifts the types of ALI metrics but the fundamental single and multi-factor analysis will, most likely, remain viable. The nature of a TOC data gathering, algorithm development, and tool may have to be reengineered to ensure customer acceptance and TOC problem relevance. Specifically, the following areas will need addressing:

- What are the salient TOC assessment goals and objectives?
- What are the ALI metrics most relevant to TOC assessment?
- What TOC ALI human interaction interfaces would be most useful to users?

Qualification and acceptance metrics—We will continue to investigate how ALI metrics (or derivatives) might be viable for also monitoring, controlling, predicting, and maximizing success of system qualification testing. Expanding and defining metrics and methods relative to predicting and analyzing program qualification and acceptance test success remains a goal.

References

- Arena, M.V.L., Robert, S., Murray, S.E., Younossi, O. (2006). Historical cost growth of completed weapon system programs. *RAND Project Air Force, TR-343*.
- Buchanan, C., & Jungbluth, C. (2010). *Leading indicators for OPEVAL success*. Unpublished Capstone Project Report, Naval Postgraduate School, Monterey, CA.
- Denq, T., Hein, G., & Gilliard, J. (2009). *Applied Leading Indicator (ALI) development methodology, appendix A*.
- INCOSE. (1995). *Metrics guidebook for integrated systems and product development*. Seattle, WA: Author.
- INCOSE. (1998). *Systems engineering measurement primer* (Ver. 1.0). Seattle, WA: Author.



INCOSE. (2010). *Systems engineering handbook, a guide for system life cycle processes and activities v.3.2* (No. INCOSE-TP-2003-002-03.2).

NASA. (2007). *Systems engineering handbook*.

NAVY. (2004). *Naval systems engineering guide*.

Rhodes, D.H., Valerdi, R., & Roedler, G.J. (2009). Systems engineering leading indicators for assessing program and technical effectiveness. *Systems Engineering*, 12(1), 15.

Roedler, G.J. (2005). *Technical measurement, a collaborative project of PSM, INCOSE, and Industry* (ICOSE Technical Report INCOSE-TP-2003-020-01).

Roedler, G.J., & Rhodes, D.H. (2007). *Systems engineering leading indicators guide* (Ver 1.0).



Funding for Life: When to Spend the Acquisition Pot

Kate Gill—Kate Gill joined the Defence Science and Technology Laboratory (Dstl) in 2005; Dstl is part of the UK Ministry of Defence. Gill has had a career in engineering management, including various posts in commercial aviation and defence companies. She has an Aeronautical Engineering degree from Southampton University, UK, which was followed by a technical apprenticeship at BAE Systems (Warton, Lancashire). This role included structural assessment, stores release, trials planning and flight line operations. Subsequent roles have included technical lead on weapons and navigation systems for the design, build, test and commissioning of corvette class ships for the Middle East (Vosper Thornycroft, Southampton) and trials design, implementation and Military Aircraft Release (MAR) authority for weapons trials (QinetiQ, Boscombe Down). Interspersed with the military domain were periods of employment with commercial aviation companies (Meggitt Avionics, Penny and Giles Aerospace). These included CAA and FAA signatory responsibilities for software and hardware development. She has an MBA and is currently studying towards an Engineering Doctorate in Systems (Eng.D.) with Bristol and Bath Universities, based on research on systems-related issues affecting UK Defence. Gill would like to acknowledge support and guidance from Prof. Patrick Godfrey and Prof. Jan Noyes, both from the University of Bristol, UK.

Contact Details: Defence Science and Technology Laboratory (Dstl), Naval Systems Department, Maritime Systems Engineering Group, Portsmouth Hill Road, Portsmouth, PO17 6AD, England, kmgill@dstl.gov.uk

Kirsty Carter-Brown—Dr. Kirsty Carter-Brown, Dstl Associate Fellow, Naval Systems Department, Defence Science and Technology Laboratory (Dstl); Dstl is part of the UK Ministry of Defence. Dr. Kirsty Carter-Brown is a Dstl Associate Fellow and joined Naval Systems Department (NSD) in November 2005, following 10 years as a published academic researcher in Biological Sciences and 2 years as a Dstl Knowledge Agent. She has been involved in a variety of programmes including research on the management and development of Ministry of Defence (MoD) acquisition processes, enterprise and benefits modelling and providing advice for the longer-term technical planning for the Modular Open Systems Architecture (MOSA) programme, the Coalition Warrior Interoperability Demonstration (CWID) and the Future Mine Countermeasures Capability (FMCMC). She is currently seconded to the Systems Engineering Integration Group (SEIG), which is part of DE&S at Abbey Wood, Bristol, UK.

Contact Details: Defence Science and Technology Laboratory (Dstl), Naval Systems Department, Maritime Systems Engineering Group, Portsmouth Hill Road, Portsmouth, PO17 6AD, England, kcbrown@Dstl.gov.uk

Abstract

The UK Ministry of Defence (MoD) spends £2.6 billion per year (MoD, 2008, Section 1.2; 2007, p. 2) on research and development (R&D) (“Defense Technology,” 2009, p. 22). The figure is technically correct, but it conceals more than it reveals. Of the total, the MoD spends around £500 million (MoD, 2006, p. 8) on laboratory research and on taking what emerges from the lab in its first few steps down the long road that leads eventually to mature technology embedded in military equipment. Analysis based on UK National Audit Office (NAO) data (Stationery Office, 2006, November 24; 2008, p. 5) shows that about half of project timescale overruns are due to technology maturation occurring too late (Jordan & Dowdy, 2007, p. 16). US evidence shows that defence projects that get all their technology mature before the equivalent of Main Gate suffer only very small time and cost overruns thereafter (GAO, 2007, pp. 14-15).



The first hypothesis would be that funding for technology development occurs too late in the acquisition process, when the problems that inevitably occur have a disproportionate effect on project timescales and costs. However, is this construct based on project failings or funding process failures? To improve the outcome, could we simply improve the timing allocation of funds? This research examines the profile of funding as aligned to maturity levels of technology and system and integration readiness, and makes proposals on the improvements that could be made.

Keywords: Finance, acquisition, maturity, maturity levels, timescales, system readiness, audit, MoD; ministry of defence; defence, defense, systems engineering, Dstl; Defence Science and Technology Laboratory, National Audit Office, NAO.

When is the Best Time to Spend—Globally?

Organisations that procure products or services through contracts are routinely involved in costing and estimation efforts. Defence Acquisition is a long-term activity, and the success or failure of the endeavor can, in part, be related to the stability of the aims and contributory components.

Economic growth has been driven by globalisation over the last 30 or more years, generating networks of connections and interdependencies between the major economic powers that are unprecedented in extent and pervasiveness (DCDC, 2009). The economic landscape has evolved rapidly with the demise of centrally planned economies, such as the Soviet Union; the rise of Asian economies, particularly China, which has embraced a market aware philosophy; and the maturation of the European Union as a cohesive economic market. These changes have created a multi-faceted economic landscape that is intimately interconnected and influential.

Over the last 30 years, the global economy has grown at a trend rate⁶² of 3-4% (IMF, 2008), and output has seen a greater-than four-fold increase. Economic growth, combined with the continuing rise in the global population, will intensify the demand for natural resources, minerals, and energy. It will continue to be uneven, fluctuating over time and between regions (Goldman Sachs, 2007). This global economic context places more pressure on all dimensions of UK government funding, and with the economic recession being felt in the homeland environment, defence focus shifts from operational reach to border security. Operational reach is the defensive objective of nations in times of affluence; border security is the focus during recession.

The IMF considers there have been five global recessions in the last 30 years (IMF, 2008). Global economic recessions will happen over the next few decades, and governments are likely to respond to them with protectionist policies, designed to shield their own economies and workforces from global competition. The global recession that started in 2008 is illustrative of the likely response to, and effects of, future recessions. One outcome of a global recession is an expected increase in the incidence of poverty, which will promote grievance and dissatisfaction among those who suffer economic hardship, in turn breeding political violence, criminality, societal conflict and the destabilization of states and regions unable to cope. This type of economic crisis will produce pressures on government finances that, considered in isolation, will place downward pressure on global defence spending.

62 Trend economic growth refers to average growth of an economy over a cycle of expansion and contraction. It is a moving average. Global recession is defined as annual world growth (based on purchasing-power-parity weights) falling below 3%. By that measure, four periods since 1980 qualify: 1980-1984, 1990-1993, 1998, 2001-2002, plus 2008 onwards. The figure of 3% is a rough correlation to negative global GDP per capita growth (IMF, 2008).

In September 2009 (“Britain’s Sharp,” 2009), the UK Chancellor of the Exchequer forecasted that in April 2010 public borrowing would exceed 12% of GDP. This projection would give Britain the biggest debt build-up between 2007 and 2014 in the G7 economies. Gaining credibility will require clear targets to cut the deficit, together with detailed and transparent plans on how to achieve them. A crucial decision will be the mix of spending cuts and tax rises. In the UK, public spending represents close to 50% of GDP, according to the Treasury. In the March 2010 UK Budget, the plan proposed was to halve the level of borrowing, to 5.2% of national income, by 2013-14 (“Economic,” 2010). The chancellor stated that spending plans from 2011 will be “*very tough—the toughest for decades,*” but refused to outline how this will affect public services, except to say that the Government would “*protect spending on those frontline public services on which we all depend.*” The annual funding review effects all government departments and has inter-related impacts on the services and support that can be offered. In the US, the current focus has been on the US health systems and associated reforms. In the UK, balancing the books and the impending election process has raised the level of scrutiny and resulted in long-term plans being put on hold until the political direction and stability have been secured (Arron, Gale & Orszag, 2004).

In March 2010, UK expenditure per government department can be represented as a pie chart showing the key funded government departments (Figure 1) (“HM Treasury,” 2010). The allocation of spending to functions is largely based on the United Nations’ *Classification of the Functions of Government (COFOG)* guidelines. Other expenditures include general public services, including international services, recreation, culture and religion, public service pensions, plus spending yet to be allocated and some accounting adjustments. The Defence spending is £40 billion, which represents 5.6% of the £704 billion total.

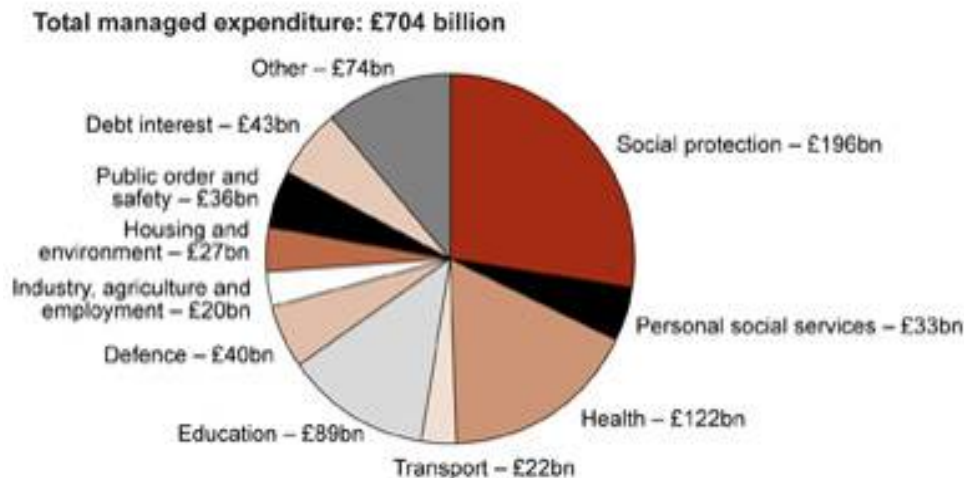


Figure 1. UK Government Spending Allocation 2009/10

Looking at the global and national data, it could be argued that “no” it is not the time to spend. Any expenditure commitments at this time will disappear into the deficit arrears, and any actual gains will be lost. However, funding streams need to be spent to ensure that the position does not regress, and so investment is required to ensure the maintenance of current position and level of capability and service.

In the current global and national climate, it is not a good time to spend. However, to maintain capability in defence and other sectors, spending has to be sustained to ensure stability and maintenance. In terms of defence projects, some projects are pending approval, some projects are committed by still under development and some projects are further down the in-service and support stage. For committed projects, the global and national situation can be considered as relevant, but not a significant issue. The national situation will determine the size of the overall budget, but how and when it is spent is an issue for the project team. Systems engineers often serve as technical points of contact throughout the entire system lifecycle (Smartt & Ferreira, 2010); as such, they are one of the critical repositories of project life knowledge, technically and financially (Valerdi, Rieff, Roedler, Wheaton & Wang, 2007). Using the natural holistic skills of a systems engineer, and applying those inherent skills to costing, has been the basis of this research.

To date, there has been significant research into through life costing (Haskins, Forsberg & Krueger, 2007; Valerdi & Miller, 2007),⁶³ and the projection of through life implications is based on existing data, extrapolated to a conclusion. This method is well documented and appealing, as it reinforces the view that we can corporately learn from experience and that costing is achieved within a defined set of rules. It, therefore, follows that if we can define the rules, then we can define the outcome (Stationery Office, 2007; 2008). This research examines the project costing profile, as related to three measures of readiness that are well defined in the systems domain. These three were selected because they represent a global viewpoint and are used in a number of multinational and multilingual programmes.

Optimum Technology, System and Integration Investment Date

Technology assessment of a complex system requires the assessment of all of its systems, sub-systems and components. Technology Readiness Levels (TRLs) can be used to determine the current component maturity. Although broadly similar, different international agencies use different definitions for TRLs; the most common definitions are those used by the Department of Defense (DoD) (DoD, 2006) and the National Aeronautics and Space Administration (NASA) (Mankins, 1995) (Table 1 and Figure 2). In Europe, the European Space Agency's (ESA) definitions of TRLs are often used as contractual baseline criteria for project acceptance (ESA, n.d.). A TRL calculator has been developed for the United States Air Force by Nolte et al. (2003). As with all assessments, this TRL calculator provides a snapshot of technology maturity at a given point in time. The Technology Program Management Model (TPMM) was developed by the United States Army by Craver et al. (2006). The TPMM is a TRL-gated high-fidelity activity model that can be used as a management tool to assist in planning, managing, and assessing technologies for successful technology transition. The model provides a core set of activities, including systems engineering and program management tasks, that are tailored to the technology development and management goals.

⁶³ Through Life Costing is a defined UK Ministry of Defence (MoD) process (Acquisition, 2009).

Table 1. Table of TRL Descriptions

National Aeronautics and Space Administration (NASA) Technology Readiness Level
1. Basic principles observed and reported
2. Technology concept and/or application formulated
3. Analytical and experimental critical function and/or characteristic proof of concept
4. Component and/or breadboard validation in laboratory environment
5. Component and/or breadboard validation in relevant environment
6. System/subsystem model or prototype demonstration in a relevant environment (ground or space)
7. System prototype demonstration in a space environment
8. Actual system completed and 'flight qualified' through test and demonstration (ground or space)
9. Actual system 'flight proven' through successful mission operations

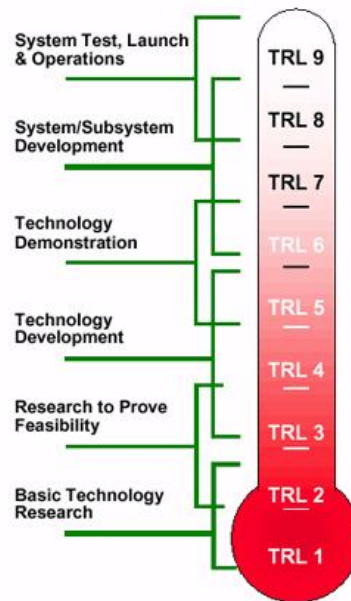


Figure 2. NASA TRL Diagram

System Readiness Levels (SRLs) ("What Are," 2009) have been developed as a project management tool to capture evidence, and assess and communicate system maturity in a consistent manner to stakeholders. In the UK, SRLs define a set of nine maturity steps from concept to in-service across a set of systems engineering disciplines. Each of the SRL steps align to key outputs from systems disciplines such as Training, Safety and Environmental, or Reliability and Maintainability. SRLs aim to take a consolidated view of the essential steps needed to properly mature and deliver a complete supportable system to the end user. SRLs are a means of analysing key outputs of an equipment acquisition project, structured in such a way as to provide an understanding of work required to mature the project. In the UK, a matrix is used to capture these key outputs and understand how they should mature over time. Projects track their maturity through the

nine SRL steps across all relevant system disciplines, and a simple self assessment tool is provided (Figure 3).

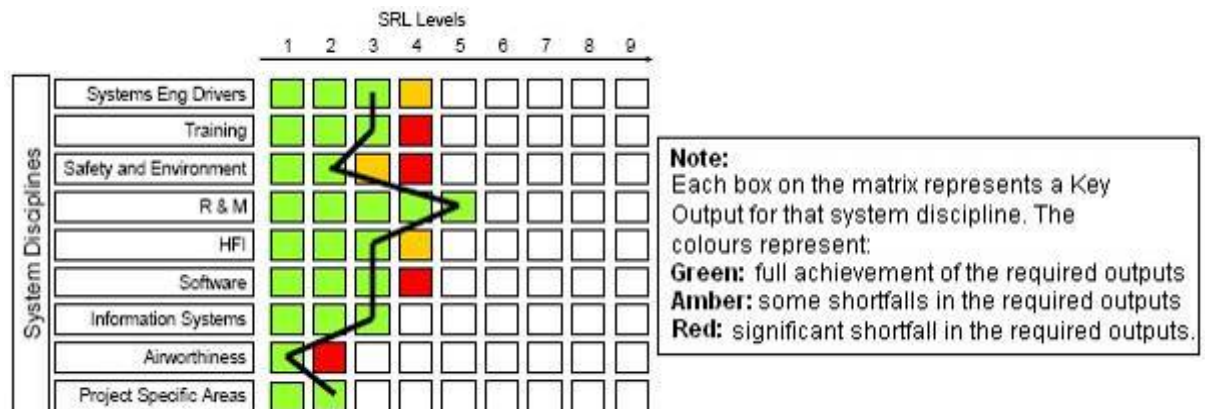


Figure 3. Outline of UK SRL Matrix with a Representative Project SRL Signature Matrix
(Applying, n.d.)

In the generic SRL matrix, each SRL level from 1 to 9 is broken down into key outputs that need to be achieved for each of the systems disciplines, including Systems Engineering Drivers, Training, Safety and Environment, Reliability and Maintainability, Human Factors Integration, Software, Information Systems, Airworthiness, Project Specific Areas. For any project, an SRL assessment produces a “signature” rather than an absolute SRL figure. The signature records the variation of maturity that has been achieved across the systems disciplines, acknowledging that not all projects mature against the systems disciplines at a consistent rate. The signature shown in Figure 3 would be typical of a project in the Assessment phase (i.e., post Initial Gate where the target is to achieve SRL1), but indicates insufficient maturity to proceed to Main Gate (the target being SRL4).

Integration levels are the next step. When brought together with TRLs and SRLs, Integration Readiness Levels (IRLs) provide a means of progressively measuring project maturity at technology, component, sub system and whole system levels. Maturity is measured in terms of cost, schedule and risk and should also take into consideration the impact of the overall system on the systems with which it is to operate, i.e., in a “system of systems” context. IRLs are used at a prime contractors and systems integrator level, and aim to support successful transition of technologies to and between acquisition partners. IRLs are used to support strategic, business and political progression, such as funding considerations (national or private), economic offsets and inter-government partnerships.

We can consider the three (TRL, SRL, IRL) measurements as interrelated but independent sets of assessment, and use these assessments to build up a visual and numerical representation of maturity. Using the Red, Amber, Green (RAG) notation, this representation resembles a wall of bricks (Figure 4) and is, therefore, termed a wall matrix view. The lowest, and foundation levels, are the TRLs, and the number of layers relates to the number of key components. The next tranche are the SRLs that cover the nine key categories. The top, capping, level is the IRL, which can also be referred to as the System of System (SoS) integration level. Any project, of any complexity, can be evaluated using the wall matrix construct. As with any assessment, the wall matrix is linked to a specific time in the life of the project, and so represents a snapshot of maturity. Relating these “time now” wall matrix assessments to the costing lifecycle was the next step of the research.



Figure 4. TRL, SRL, IRL Wall Matrix Construct

Costing Lifecycle

The costing lifecycle is a method of project representation in which all costs arising from owning, operating, maintaining and disposing of a project are considered. Project costs take into account the initial cost, including capital, investment cost, development cost, unit purchase and installation cost, and future costs, including operating cost, energy cost, maintenance cost, capital replacement cost, financing cost and any resale salvage for disposal cost over the life time of a project or technology product. Any new product or project progresses through a sequence of stages (Figure 5). Figure 5 expresses the profile against an axis of sales, but the axis could equally represent investment, effort or cost.

Using the traditional costing profile, and nesting the contribution of each of the stages of the UK Defence Acquisition CADMID⁶⁴ lifecycle, a series of contributory peaks and troughs can be mapped. Using Valerdi's results alongside similar project management costing models from Kerzner (2003), a lifecycle costing profile has been constructed (Figure 6) and then roughly aligned to the UK CADMID process. Each contributory cycle has a peak, which represents the achievement of the aim of that stage of the lifecycle.

There are a number of reports that relate to early investment, resulting in lower overall project cost. Jorgensen and Sjoberg (2001) examined a portfolio of software projects, and examined error occurrences as related to overrun costs. Their results show early investment was associated with early development of maturity, and, as a result, minimum rework and lower overall project cost (Jorgensen & Molokken-Ostovold, 2004; Jorgensen & Grimstad, 2005; Jorgensen, 2009). Valerdi (2005) and Valerdi et al. (2007) constructed a cost modeling tool that they validated against industrial projects. Using this evidence with the nested costing lifecycle, it can be concluded that contribution of funding

⁶⁴ CADMID means Concept, Assessment, Demonstration, Manufacture, In-Service and Disposal.

prior to the peak accelerates that stage of the process; contribution after the peak ensures appropriate closure of that stage and contributes to the start up of the next stage.

The research then focused on assessing when contributory peak was reached, in relation to the maturity levels, with the premise that investment prior to the peak was constructive up to the peak point, at which point it was better to invest in the next ascending stage of the process.

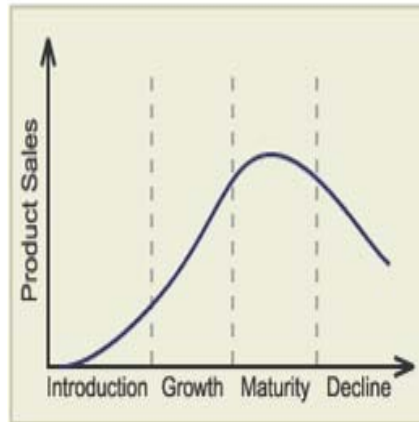


Figure 5. Product Lifecycle Diagram
(Gorhcels, 2000)

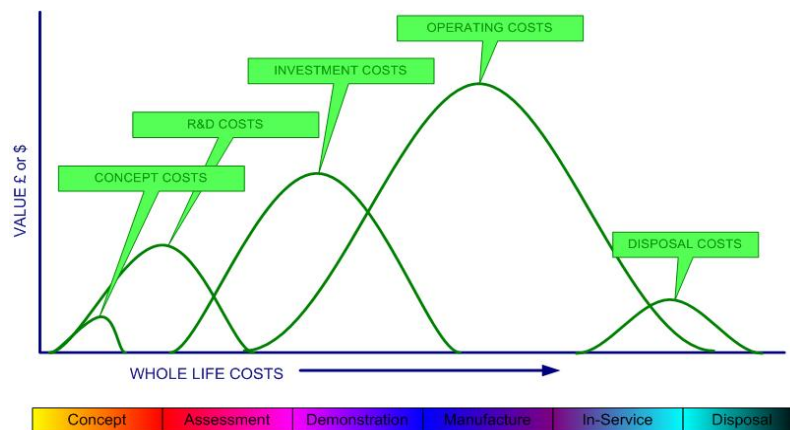


Figure 6. Lifecycle Costing with Key Contributing Stages Mapped Against CADMID

Costing Case Studies

Using historical cost and progress data, a number of projects have been examined. The cost data has been sourced from UK National Audit Office (NAO) reports (Stationery Office, 2006, November 24, p. 14; 2008, p. 5; 2006, p. 57; 2007a, p. 4; 2007b, p. 6,313; 2003) to minimize supplier interpretation, and selective accounting. The project data has also been sourced from NAO reports alongside other open source publications.

Unlike other studies, access to data was not the issue; the data was in abundance. The real issue was in consistent interpretation from the differing types of source document and in the maturity assessments, which were based on technical reports and publications. For the purposes of this openly published paper, the case studies have been deliberately censured. All the source data is available openly, but some has been interpreted in the context of information that is privileged to the UK MoD.

During the period August 2009 and March 2010, twelve projects were researched for source data. The research related to gaining published material that could be assessed against the costing and the three maturity criteria. The aim was to provide source data that could be validated by the authors and by a technical judgment panel of engineers and scientist proficient with the UK acquisition lifecycles and level assessments. The panel was sourced from colleagues with Dstl, who have familiarity with the assessment process and minimal knowledge or involvement in the selected projects. The projects were assessed against their known timeline, the theoretical costing lifecycle, published costs and associated panel TRL, SRL, IRL assessments. Each project was assessed at various points in its lifecycle. The number of assessment points varied from two to twelve. Larger, long-term projects were replete with published data and so were assessed at multiple points; for some smaller projects, data was only available at two points. For each project, a consistent selection of components was assessed for TRL. The TRLs, the nine aspects of SRL and the capping IRL were then discussed using the source data and assessed on a wall matrix by the panel.

Once a timeline position and an associated “RAG” wall matrix had been established, the average TRL, SRL, IRL score was calculated and plotted against the known cost curve of the project. Some plot points were on an ascending line, e.g., prior to peak, and some were on the declining line. In total, twenty eight wall matrix plots were constructed and associated with a costing timeline. Some of the projects were reported on for three consecutive years of NAO assessment so that a robust element of progression could be established. After ratification and review, a consolidated plot was constructed (Figure 7).

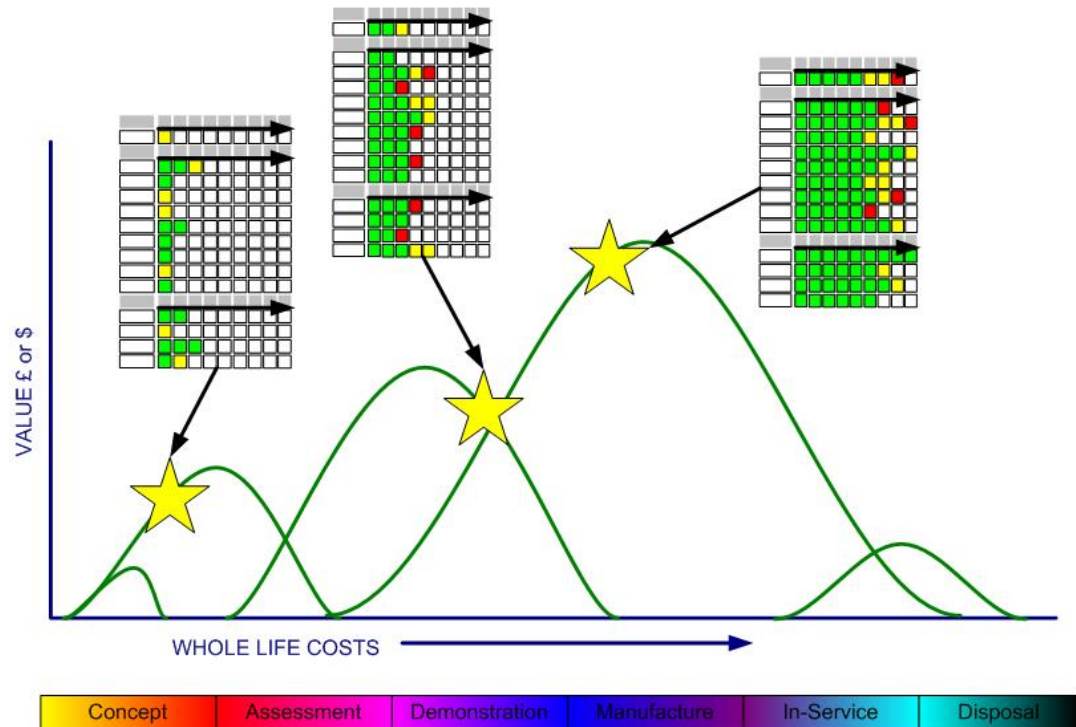


Figure 7. Consolidated Plot of Acquisition Lifecycle Stages versus Readiness Levels Matrices

Overview, Conclusions and Further Work

After reviewing and consolidating the data, a table was produced comparing the acquisition lifecycle phases of CADMID against the set of points of significance for the three readiness levels, extracting for the wall matrix (Table 2). From the data, the optimum investment date could not be established, but a limit point was found. Investments in a particular stage added benefit to a TRL, SRL, IRL limit point; investment after that point was required to ensure project progression, but did not represent the peak exchange.

Table 2. Consolidated Wall Matrix and Timeline Table of Results

	TRL point of significance	SRL point of significance	IRL point of significance
CONCEPT	< 3.2	< 2.4	N/A
ASSESSMENT	< 5.6	< 6.5	< 1.8
DEMONSTRATION	< 7.6	< 7.8	< 4.4
MANUFACTURE	< 8.1	< 8.4	< 8.2
IN-SERVICE	N/A	N/A	< 9.0

There was a large variation in the sets of TRL, SRL and IRL data that will be examined in further studies. The source of variation could be related to source data error, interpretation or factors that made the specific project different to its peer group, e.g.,

groundbreaking atomic physics as opposed to next generation vehicle development. Expanding the analysis to a wider pool to projects would refine the technique and increase the robustness of the “points of significance” as well as reduce the statistical effects. The use of one technical judgment panel was convenient for this stage of the research but could introduce experimental errors over time. Familiarity and prior knowledge of previous results started to bias conversations on some of the latter assessments of larger projects.

Looking at the global and national economic situation, the competition for government funding is becoming more intense, and departments such as the MoD are under increasing pressure to deliver optimum results against investment. This investment is not just what to invest in, but also when to invest. The extant UK MoD assessment criteria include TRL, SRL and IRL decision points, so this data is already available. By mapping the optimum investment point against the TRL, SRL, and IRL, a wall matrix could guide the diversion of funds into projects up to key points, which would result overall in maximizing the outcome benefit to UK Defence.

References

- Acquisition operating framework* (Ver. 2.0) (2009, December 19). Retrieved January 18, 2010, from <http://www.aof.mod.uk/aofcontent/tactical/tlcm/content/introductiontotlcm.htm>
- Applying system readiness levels*. (n.d.). Retrieved from http://www.aof.mod.uk/aofcontent/tactical/techman/content/srl_applying.htm
- Arron, H.J., Gale, W.G., & Orszag, P.R. (2004). Meeting the revenue challenge. In A.M. Ravlin & I.V. Swahill (Eds.), *Restoring fiscal sanity: How to balance the budget* (pp. 111-127). Washington, DC: The Brookings Institution.
- Britain's sharp budget choices. (2009, September 24). *The Economist*. Retrieved from <http://www.economist.com>
- Craver, J.T., et al. (2006). *Technology program management model*. Paper presented at the NDIA Systems Engineering Conference, Army Space and Missile Defense Command Technical Center.
- The defence technology plan. (2009, June). In *Proceedings of the RUSI Defence Systems Conference*.
- Development, Concepts and Doctrine Centre (DCDC). (2009, July 16). Dimensions—research areas. In *DCDC Strategic Trends Programme: Global Strategic Trends* (pp. 87-158) (4th ed.). UK: Crown Copyright.
- DoD. (2006, July 7). Technology readiness levels in the Department of Defense (DOD). In *Defense Acquisition Guidebook* (Ver. 1.6). Washington, DC: Author.
- Economic and financial indicators. (2010, March 18). *The Economist*. Retrieved from <http://222.economist.com>.
- European Space Agency (ESA). (n.d.) Retrieved from <http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=37710>
- GAO. (2007, March). *Defense acquisitions: Assessments of selected weapon programs* (GAO-07-406SP). Washington, DC: Author.
- Goldman Sachs Group. (2007, November). Beyond the BRICS: A look at the “next 11.” In *BRICs and Beyond* (pp. 159-164). New York: Author.
- Gorhcels, L. (2000). *The product manager's handbook: The complete e-management resource* (2nd ed.). London: McGraw-Hill.
- Haskins, C., Forsberg, K., & Krueger, M.J. (2007). *System engineering handbook: A guide to systems lifecycle processes and activities*.

- HM Treasury 2010-11 near cash projections. (2010, March 18). *The Economist*. Retrieved from <http://economist.com>
- International Monetary Fund (IMF). (2008, October). *World economic outlook (WEO) 2008: Financial stress, downturns, and recoveries* (pp. 3-44). Retrieved October 16, 2008, from <http://www.imf.org/external/pubs/ft/weo/2008/02/pdf/text.pdf>
- Jordan, G., & Dowdy, J. (2007, October). Acquisition and technology: The problem that won't go away. *RUSI Defence Systems*, 10(2).
- Jorgensen, M. (2009). How to avoid selecting bids based on overoptimistic cost estimates. *IEEE Software*, 79-84.
- Jorgensen, M., & Grimstad, S. (2005). Over-optimism in software development projects: "The winner's curse." In *Proceedings of the 15th International Conference on Electronics, Communications and Computers (CONIELECOMP 2005)*.
- Jorgensen, M., & Molokken-Ostovold, K. (2004). Reasons for software effort estimation error: Impact of respondent role, information collection approach, and data analysis method. *IEEE Transactions on Software Engineering*, 30(12), 993-1007.
- Jorgensen, M., & Sjøberg, D.I.K. (2001). Impact of effort estimates on software project work. *Information and Software Technology*, 43, 939-948.
- Kerzner, H.P. (2003). *Project management: A systems approach to planning, scheduling, and controlling*. Hoboken, NJ: Stevens Institute of Technology.
- Mankins. (1995). *Technology readiness levels in the National Aeronautics and Space Administration (NASA): A white paper*. Washington, DC: NASA.
- MoD. (2006, September). *Maximising benefit from defence research*. Working paper.
- MoD. (2007, October). *Maximising defence capability through R&D*. Report.
- MoD. (2008, September). *Technology partnership in defence*. Working paper.
- Nolte, W.L., et al. (2003). *Technology readiness level calculator*. Paper presented at the NDIA Systems Engineering Conference, Air Force Research Laboratory.
- Smartt, C., & Ferreira, S. (2010, March 17-19). Leveraging psychology and software engineering research to develop bid, proposal, and cost estimation heuristics for systems engineering (Paper #1569270565). In *Proceedings of CSER 2010, 8th Conference on Systems Engineering Research*. Hoboken, NJ: Stevens Institute of Technology.
- The Stationery Office. (2003). Major Projects Report 2003 (43rd Report). HC 383, Session 2003-2004. London: National Audit Office, Comptroller and Auditor General.
- The Stationery Office. (2006). *Ministry of Defence: Major projects report—project summary sheets*. HC23-1 Session 2006-2006. London: National Audit Office, Comptroller and Auditor General.
- The Stationery Office. (2006, November 24). *Ministry of Defence: Major projects report 2006*. HC23-1, Session 2006-2007. London: National Audit Office, Comptroller and Auditor General.
- The Stationery Office. (2007a). *Comptroller and Auditor General annual report* (Vol. 1). London: National Audit Office, Comptroller and Auditor General.
- The Stationery Office. (2007b). *Comptroller and Auditor General annual report* (Vol. 2). London: National Audit Office, Comptroller and Auditor General.
- The Stationery Office. (2008, July 22). *Ministry of Defence: Major projects report 2007*. HC433. London: National Audit Office, Comptroller and Auditor General.
- The Stationery Office, National Audit Office (NAO). (2007). *Ministry of Defence: Major projects report 2007*. Retrieved November 30, 2007, from http://www.nao.org.uk/publications/0708/ministry_of_defence_major_pro.aspx
- The Stationery Office, National Audit Office (NAO). (2008). *Ministry of Defence: Major projects report 2008*. Retrieved December 18, 2007, from http://www.nao.org.uk/publications/0809/mod_major_projects_report_2008.aspx

- Valerdi, R. (2005). *The constructive systems engineering cost model (COSYSMO)*. Unpublished doctoral dissertation, University of Southern California.
- Valerdi, R., & Miller, C. (2007). From research to reality: Making COSYSMO a trusted estimation tool in your organization. In *Proceedings of International Council on Systems Engineering (INCOSE) 2007 Symposium*. San Diego, CA.
- Valerdi, R., Rieff, J.E., Roedler, G.J., Wheaton, M.J., & Wang, G. (2007). Lessons learned from the industrial validation of COSYSMO. In *Proceedings of the International Council on Systems Engineering (INCOSE) 2007 Symposium*. San Diego, CA.
- What are system readiness levels (SRLs)? (2009, December 19). *Acquisition operating framework: Technology Management*. Retrieved from http://www.aof.mod.uk/aofcontent/tactical/techman/content/srl_whatarethey.htm



System Capability Satisficing in Defense Acquisition via Component Importance Measures

Brian Sauser—Brian J. Sauser received his BS in Agriculture Development from Texas A&M University, MS in Bioresource Engineering from Rutgers, The State University of New Jersey, and PhD in Technology Management from Stevens Institute of Technology. He is currently an Assistant Professor at Stevens Institute of Technology in the School of Systems and Enterprises.

Brian J. Sauser

Stevens Institute of Technology
School of Systems and Enterprises
One Castle Point on Hudson
Hoboken, NJ 07030
Tel: (201) 216-8589
Fax: (201) 216-5541
bsauser@stevens.edu

Weiping Tan—Weiping Tan received his BE in Automation (first class honors) in 2006 from Beijing Institute of Technology, and received his ME in Engineering Management from Stevens Institute of Technology in 2009. He is currently pursuing his PhD in Engineering Management at Stevens Institute of Technology in the School of Systems and Enterprises.

Weiping Tan
Stevens Institute of Technology
School of Systems and Enterprises
One Castle Point on Hudson
Hoboken, NJ 07030
Tel: (201) 216-8118
Fax: (201) 216-5541
wtan@stevens.edu

Jose E. Ramirex Marquez—Jose E. Ramirez-Marquez received his BS in Actuarial Science from Universidad Nacional Autónoma de México, MS in Statistics and MS and PhD in Industrial and Systems Engineering from Rutgers, The State University of New Jersey. He is currently an Assistant Professor at Stevens Institute of Technology in the School of Systems and Enterprises.

Jose E. Ramirez-Marquez
Stevens Institute of Technology
School of Systems and Enterprises
One Castle Point on Hudson
Hoboken, NJ 07030
Tel: (201) 216-8003
Fax: (201) 216-5541
jmarquez@stevens.edu



Romulo B. Magnaye—Romulo B. Magnaye received his BS in Mining Engineering and MBA from the University of the Philippines and a postgraduate research Diploma in Mineral Technology from Camborne School of Mines in England. He is currently a Robert Crooks Stanley fellow pursuing his PhD in Engineering Management at Stevens Institute of Technology in the School of Systems and Enterprises. Romulo B. Magnaye

Stevens Institute of Technology
School of Systems and Enterprises
One Castle Point on Hudson
Hoboken, NJ 07030
Fax: (201) 216-5541
rmagnaye@stevens.edu

David Nowicki—David Nowicki received his BS and MS in Industrial Engineering and Operations Research from the University of Wisconsin and Virginia Tech respectively, and PhD in Industrial and Systems Engineering from the University of Wisconsin. He is currently an Associate Professor at Stevens Institute of Technology in the School of Systems and Enterprises.

David Nowicki
Stevens Institute of Technology
School of Systems and Enterprises
One Castle Point on Hudson
Hoboken, NJ 07030
Fax: (201) 216-5541
rmagnaye@stevens.edu

Abhijit Deshmukh—Abhijit Deshmukh received his PhD from Purdue University. He is currently a Professor at Texas A&M University in the Department of Industrial and Systems Engineering.

Abhijit Deshmukh
Texas A&M University
Department of Industrial and Systems Engineering
College Station, TX
deshmukh@tamu.edu

Abstract

With support from the Naval Postgraduate School and government/industry partnerships, the Systems Development & Maturity Laboratory (SD&ML) at Stevens Institute of Technology has successfully developed a systems maturity measure (i.e., System Readiness Level [SRL]) and supporting optimization models for inclusion in a Systems Earned Readiness Management methodology. We now believe it is time to spiral back to the beginning of the original developments of the SRL to enhance fundamental capabilities of assessing system maturity in order to address some recurring issues to its application. That is, systems have variants in their physical architecture that realize certain functionality and capability by which trade-off decisions are made to find a satisficing solution for a deployable system. This paper enhances previously developed methodologies by addressing this fundamental question, “What are the trades-off in functionality, capability, cost, schedule, and maturity that will allow the deployment of a less-than-fully mature system that can still satisfy specific needs of the warfighter?” To answer this question, we formulate a capability-specific SRL and use multi-dimensional component importance analysis to identify which components of the system should receive the most application of resources when they are constrained.



Introduction

With the support of the Naval Postgraduate School (NPS) and government/industry partnerships, the Systems Development & Maturity Laboratory (SD&ML) at Stevens Institute of Technology has successfully:

5. Developed a methodology for determining a system's maturity using the System Readiness Level (SRL) scale (Sausser et al., 2008);
6. Used SRL to formulate two optimization models—SRLmax (Sausser, Ramirez-Marquez, et al., 2008) and SCODmin (Magnaye et al., 2010)—for predicting cost, schedule, and maturity performance, and
7. Proposed a methodology that combines items a and b into an approach called Systems Earned Readiness Management (SERM) (Magnaye et al., 2009).

During the research that has fostered these developments, the SD&ML has maintained a spiral development approach where we have worked closely with industry and government to refine and implement our research in order to maintain its relevance and rigor. The outputs of our research to date have focused on the analysis of systems that were meant to deliver a single capability. We now believe it is time to spiral back to the beginning of the life cycle (see Figure 1) to enhance fundamental capabilities of the SRL in order to address a situation where the system under development is intended to be multi-functional and is needed sooner rather than later by the warfighter.

Such systems are becoming the norm given the urgency of the armed conflicts that we are currently fighting in Iraq, Afghanistan, and against terrorism. However, the desire to deliver capabilities immediately has to be moderated by the reality that resources are tight, such that where possible, we must deliver systems that can accomplish multiple things in the field. Thus, we are confronted with the situation where complexity is increased by multi-functionality while the development time available is shorter.



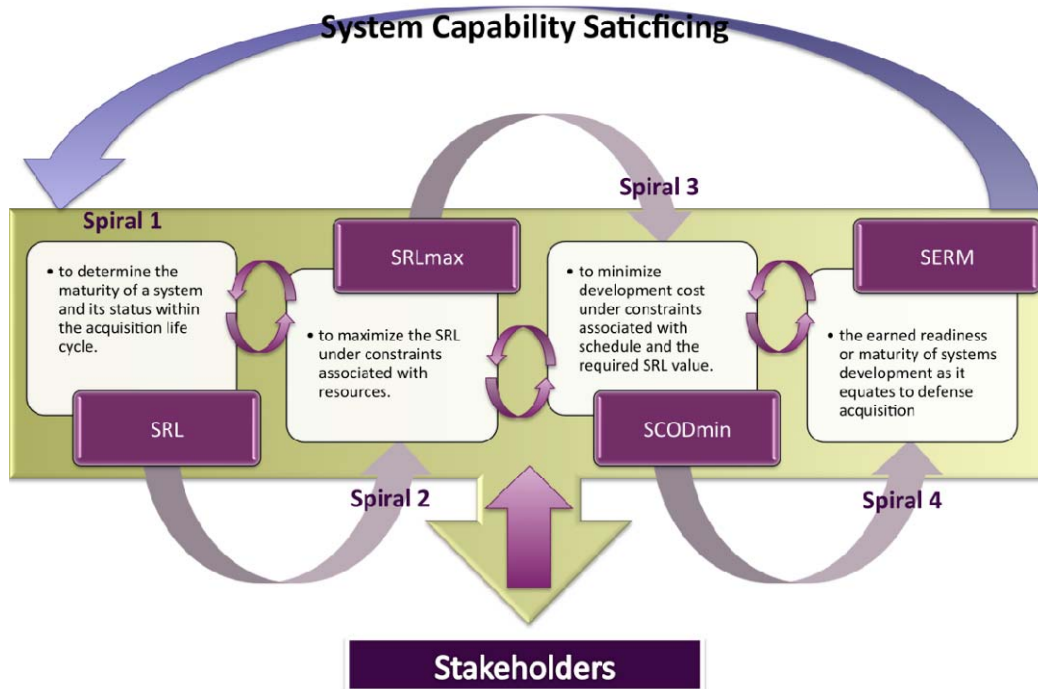


Figure 1. Research Spiral Development Plan

Since these systems are to provide multiple functionalities and capabilities, their complexity surpasses that of systems providing only single functionality with single capability. In order to secure the intended capabilities, the US government mandates that the Key Performance Parameters (KPP) must be specified in the Capability Development Document (CDD) and Capability Production Document (CPD), and be verified by testing and evaluation or analysis (DAU, 2009). According to the DAU Manual, Key Performance Parameters (KPP) are “those attributes or characteristics of a system that are considered critical or essential to the development of an effective military capability and those attributes that make a significant contribution to the characteristics of the future joint force as defined in the Capstone Concept for Joint Operations (CCJO).”

However, development of methodologies that can accurately predict the ability of systems to satisfy KPP while development is still ongoing poses challenges to the acquisition community. Volkert (2009) proposed an approach that used System Readiness Level (SRL) as the indicator of the level of capability that is being realized. It answered the question of how to predict the achieved ability of a system while its development is underway. The objective of doing so is to monitor the system developmental progress and identify issues in a timely manner should there be any gap (e.g., schedule, cost, etc.) between the preset plan and the accomplishments. The information gathered from his approach is very important because when it is coupled with further analysis, corresponding decisions can be made and measures can be taken so as to prevent the issues from getting worse. However, although Volkert's approach can provide timely data about the progress of the system, it does not address the question of how to solve development problems (should there be any), and thus is unable to prescribe methodologies for prioritizing resources allocation. Specifically, as problems arise or are anticipated, the program manager must be able to determine which of the system's components should receive resources based on their importance in achieving the capability in question.

This paper proposes such a methodology based on component importance analysis. We use the term capability to represent both the “capability” and the “physical structure of technology packages whose combination enables the capability. The paper reviews Volkert’s approach for measuring the achieved KPPs in a system under development, and then proceeds by proposing three component importance measures. A system with notional data is presented to illustrate the application of the proposed IMs. The paper ends with the conclusion the exploration of future research.

Methodology for Measuring Achieved Capability

In the procurement and management of the Mission Packages (MPs) for a system, the designated program office and manager, such as the PMS420 for the Littoral Combat Ship (LCS) program, requires insight into the progress of the Development Program Offices’ (DPOs) constituent mission systems and knowledge about where they stand in terms of providing anticipated performance, especially the KPPs of the system. These insights are critical for requisite oversight and for management of development risks. However, the issue is how program managers accurately can predict the ability of the system to satisfy KPPs while development and integration are proceeding.

Previously, DPOs were able to use Technical Performance Measures (e.g., Technology Readiness Levels [TRL]) to monitor the developmental status of specific technologies. With the development of complex multi-capability systems, such as the LCS, understanding the status of technologies are no longer sufficient for managers to gain the requisite level of knowledge on the extent to which the KPPs, as designated by the Capability Development Document, can be accomplished for the designated system. Volkert (2009) has pointed out the compounding reasons:

1. The capabilities from the individual constituent mission systems are often being modified or utilized in manners different than that specified in their original requirement set. Thus, their known/predicted performance may be different when used in a MP SoS.
2. The constituent mission systems (capabilities) being developed by the DPOs are, in some cases, still maturing. This impacts the ability to determine KPP performance in two ways;
 - a. It drives an incremental fielding of capabilities by PMS 420, meaning the solution set for accomplishing (full or partially) a KPP will vary over time.
 - b. The capability delivered by the DPO may not provide the amount of performance anticipated/predicted by PMS 420 due to developmental challenges within the DPOs program.
3. The combination of capabilities available to choose from means that the usage and contribution of an individual capability to the performance of a KPP can vary depending upon the operational employment of the system within a SoS.

Therefore, for predicting the achieved proportional capability in a complex system, Volkert proposed an approach. Here we re-write his formula with minor changes:

$$T_{C(1,2,\dots)n} = \omega_n * \alpha_{C(1,2,\dots)}$$



Where $\alpha_{C(1,2,...)}$ represents the maximum level of performance capability the combination of technology packages (1, 2, ...) is expected to meet/provide. ω_n represents the weighting factor representing the proportional level of performance expected at the maturity stage of n . $T_{C(1,2,...)n}$ thus represents the achieved performance level of the capability which can contribute to the satisfaction of the KPP. The value of α would be expressed in the units of performance defined by the KPP while ω would be unit less.

For ω_n , Volkert suggested the use of the System Readiness Level (SRL) for the capability at that time. In order to differentiate this with the original SRL definition that is designed for assessing the development maturity of the whole system, we introduce the new notion of a Capability System Readiness Level (SRL_C) to measure ω_n which represents the readiness of the Capability comprised by a specific combination of technologies and the integrations among them. For simplicity, for the rest of this paper, whenever SRL is mentioned, it means the SRL_C.

Mathematically, the procedure for calculating the SRL_C is as follows (assuming the subset of n technologies from within the system, which will have to be integrated to deliver a capability C):

- a. Normalize the [1, 9] scale original TRLs and IRLs into (0,1) scale, and denote them by matrices:

$$\mathbf{TRL} = \begin{bmatrix} TRL_1 \\ TRL_2 \\ \dots \\ TRL_n \end{bmatrix} \xrightarrow{\text{Normalize}} \mathbf{TRL}' = \frac{\mathbf{TRL}}{9} = \begin{bmatrix} TRL'_1 \\ TRL'_2 \\ \dots \\ TRL'_n \end{bmatrix};$$

$$\mathbf{IRL} = \begin{bmatrix} IRL_{11} & IRL_{12} & \dots & IRL_{1n} \\ IRL_{21} & IRL_{22} & \dots & IRL_{2n} \\ \dots & \dots & \dots & \dots \\ IRL_{n1} & IRL_{n2} & \dots & IRL_{nn} \end{bmatrix} \xrightarrow{\text{Normalize}} \mathbf{IRL}' = \frac{\mathbf{IRL}}{9} = \begin{bmatrix} IRL'_{11} & IRL'_{12} & \dots & IRL'_{1n} \\ IRL'_{21} & IRL'_{22} & \dots & IRL'_{2n} \\ \dots & \dots & \dots & \dots \\ IRL'_{n1} & IRL'_{n2} & \dots & IRL'_{nn} \end{bmatrix}$$

Where $IRL_{ij} = IRL_{ji}$. When there is no integration between two technologies, an original IRL value of 0 is assigned; for integration of a technology to itself, an IRL value of 9 is used, that is original $IRL_{ii} = 9$.

- b. ITRL matrix is the product of TRL and IRL matrices:

$$\mathbf{ITRL} = \mathbf{Norm} \times \mathbf{IRL}' \times \mathbf{TRL}'$$

That is,

$$\begin{aligned}
\begin{bmatrix} ITRL_1 \\ ITRL_2 \\ \dots \\ ITRL_n \end{bmatrix} &= \begin{bmatrix} 1/m_1 & 0 & \dots & 0 \\ 0 & 1/m_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1/m_n \end{bmatrix} \times \begin{bmatrix} IRL'_{11} & IRL'_{12} & \dots & IRL'_{1n} \\ IRL'_{21} & IRL'_{22} & \dots & IRL'_{2n} \\ \dots & \dots & \dots & \dots \\ IRL'_{n1} & IRL'_{n2} & \dots & IRL'_{nn} \end{bmatrix} \times \begin{bmatrix} TRL'_1 \\ TRL'_2 \\ \dots \\ TRL'_n \end{bmatrix} \\
&= \begin{bmatrix} (IRL'_{11} TRL'_1 + IRL'_{12} TRL'_2 + \dots + IRL'_{1n} TRL'_n) / m_1 \\ (IRL'_{21} TRL'_1 + IRL'_{22} TRL'_2 + \dots + IRL'_{2n} TRL'_n) / m_2 \\ \dots \\ (IRL'_{n1} TRL'_1 + IRL'_{n2} TRL'_2 + \dots + IRL'_{nn} TRL'_n) / m_n \end{bmatrix}
\end{aligned}$$

Where m_i is the number of integrations of technology i with itself and all other technologies, and **Norm** is to normalize the SRL_i from $(0, m_i)$ scale to $(0, 1)$ scale for consistency; i.e., **Norm** = $diag[1/m_1, 1/m_2, \dots, 1/m_n]$.

a. SRL is the average of all ITRLs:

$$SRL = \frac{ITRL_1 + ITRL_2 + \dots + ITRL_n}{n} = \frac{\sum_{i=1}^n ITRL_i}{n} .$$

See Sauser et al. (2008; 2010) for a more detailed description of how to calculate and apply the SRL.

Component Importance Measures

A system has variants in its physical architecture that realize certain functionality and capability. A systems engineer or acquisition manager would make trade-off decisions to find a satisficing solution for a deployable system. Thus, in order to satisfy Key Performance Parameter during the development of the system, this paper proposes to perform component importance analysis by introducing three Importance Measures (IMs) for System Capability Satisficing (SCS) with respect to: TRL/IRL, cost, and labor-hours.

SCS with Respect to TRL/IRL (I^P)

The IM of TRL/IRL evaluates the impact of a change in the development maturity of an component (i.e. technology or integration) on system development maturity. That is, IM measures the change of the SRL when the TRL or IRL of a specific component changes from its current value to a target value. For example, let $SRL(TRL, IRL)$ denote the current SRL of the system, and $SRL(TRL, IRL | TRL_i = \overline{TRL}_i)$ ($SRL(TRL, IRL | IRL_{ij} = \overline{IRL}_{ij})$) denote the resultant SRL when TRL_i (IRL_{ij}) changes to a target maturity level \overline{TRL}_i (\overline{IRL}_{ij}) and all other TRL s/ IRL s stay on current maturity values. The definition of IM with respect to TRL/IRL (I^P) is as follows:

$$\text{For } TRL, I_i^P = \frac{SRL(TRL, IRL | TRL_i = \overline{TRL_i})}{SRL(TRL, IRL)}$$

$$\text{For } IRL, I_{ij}^P = \frac{SRL(TRL, IRL | IRL_{ij} = \overline{IRL_{ij}})}{SRL(TRL, IRL)}$$

I^P implies the effect of change in the readiness level of a given component. A component for which the variation of the readiness level results in the largest variation of the system SRL has the highest importance.

SCS with Respect to Cost (I^{CT})

Zhang et al. (2007) suggests that classical component importance analysis ignores cost, and states that it is unrealistic to evaluate the importance of components without considering the cost. Hereby, for SRL component importance analysis, we propose to consider the economic factor. This is reasonable by noting that there are always situations where system developers have to make the investment decisions based on the comparison of the immediate return on the investment of dollars needed to mature components. Presumably, especially with a tight budget, developers allocate resources to the component that can result in the highest system maturity. Therefore, we propose I^{CT} as an IM that takes into account the cost for maturing components to facilitate such comparisons. Since the cost to mature different components varies and improvements in different components have different effects on SRL, the IM that takes into account the development cost serves as a baseline to compare the investment returns from different components. Let

$CT_i = CT_{\overline{TRL_i}} - CT_{TRL_i}$ denote the associated development cost for maturing TRL_i from its current readiness level to a target level $\overline{TRL_i}$, and $CT_{ij} = CT_{\overline{IRL_{ij}}} - CT_{IRL_{ij}}$ denote the associated development cost from maturing IRL_{ij} from its current readiness level to a target level $\overline{IRL_{ij}}$. The formula to calculate the I^{CT} is as follows:

$$\text{For } TRL, I_i^{CT} = \frac{\Delta SRL}{CT_i} = \frac{SRL(TRL, IRL | TRL_i = \overline{TRL_i}) - SRL(TRL, IRL)}{CT_{\overline{TRL_i}} - CT_{TRL_i}}$$

$$\text{For } IRL, I_{ij}^{CT} = \frac{\Delta SRL}{CT_{ij}} = \frac{SRL(TRL, IRL | IRL_{ij} = \overline{IRL_{ij}}) - SRL(TRL, IRL)}{CT_{\overline{IRL_{ij}}} - CT_{IRL_{ij}}}$$

I^{CT} implies the effect of the cost to mature a given component on SRL, and the component whose readiness improvement from the investment results in the largest gain of SRL has the highest importance.

SCS with Respect to Labor-Hours (I^{LH})

Besides the consideration of cost, there are other situations (e.g., when only certain labor-hours are available) where developers care more about the return on the effort needed to improve components. Therefore, we propose another importance measure (I^{LH}) that takes into account the associated labor-hours to upgrade the component readiness level in order to mature the SRL. Let $LH_i = LH_{\overline{TRL}_i} - LH_{TRL_i}$ denote the associated development labor-hours for developing TRL_i from its current status to a target level \overline{TRL}_i , and $LH_{ij} = LH_{\overline{IRL}_{ij}} - LH_{IRL_{ij}}$ denote the associated development labor-hours for developing IRL_{ij} from its current status to a target level \overline{IRL}_{ij} , then the formula for I^{LH} is as follows:

$$\text{For } TRL, I_i^{LH} = \frac{\Delta SRL}{LH_i} = \frac{SRL(TRL, IRL | TRL_i = \overline{TRL}_i) - SRL(TRL, IRL)}{LH_{\overline{TRL}_i} - LH_{TRL_i}}$$

$$\text{For } IRL, I_{ij}^{LH} = \frac{\Delta SRL}{LH_{ij}} = \frac{SRL(TRL, IRL | IRL_{ij} = \overline{IRL}_{ij}) - SRL(TRL, IRL)}{LH_{\overline{IRL}_{ij}} - LH_{IRL_{ij}}}$$

I^{LH} implies the effect of the labor-hours or effort to mature a given component on SRL. The component whose readiness improvement from the investment of labor-hours results in the largest gain of SRL has the highest importance.

Notional Example

The following example was used in Forbes et al. (2009) to illustrate the application of SRL. The system is designed to perform six capabilities. For the illustration of the proposed methodology in this paper, it is assumed that the *mine-detection capability* that is enabled by the combination of the shaded components is the KPP of interest. This capability has six components with six integrations among them, and the corresponding TRLs and IRLs are shown in Figure 2.

The current capability SRL for the Mine-Detection is 0.622. According to the definition of SRL (Magnaye et al., 2009), this value indicates that the capability is undergoing the Engineering and Manufacturing Development phase. During this phase, the major assignments are to develop system capability or (increments thereof), reduce integration and manufacturing risk, ensure operational supportability, minimize logistics footprint, implement human systems integration, design for production, ensure affordability and protection of critical program information, and demonstrate system integration, interoperability, safety and utility.

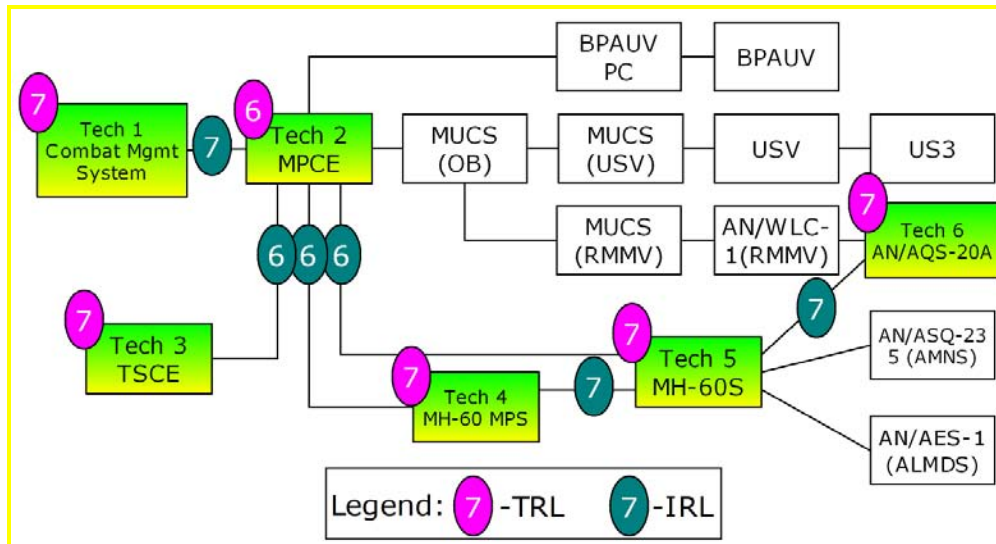


Figure 2. Diagram of a System with Components Shaded for the KPP

Since we are proposing to take into account the resource consumption (cost and labor-hour) in the component importance evaluation, Tables 1 and 2 show these values for maturing the components (i.e., TRL and IRL) of the capability of interest. The cost is in thousands of dollars (\$1,000), and the effort is in labor-hours. For example, it requires 599 hours of effort and \$980,000 to move Technology 1 from level 7 to level 8. It is the obligation of the program manager to obtain these estimates of resource consumption in reality. To mature the whole capability, the estimated cost and effort equal \$17,141,000 and 10,976 of labor-hours, respectively.

Table 1. Resource Consumption for IRL Upgrade

	Tech 1		Tech 2		Tech 3		Tech 4		Tech 5		Tech 6	
TRL	Cost	Time	Cost	Time	Cost	Time	Cost	Time	Cost	Time	Cost	Time
1	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
2	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
3	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
4	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
5	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
6	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
7	\$0	0	\$579	453	\$0	0	\$0	0	\$0	0	\$0	0
8	\$980	599	\$157	177	\$973	541	\$459	154	\$443	551	\$410	580
9	\$820	290	\$918	267	\$404	582	\$592	341	\$490	304	\$871	358
SSum	\$1,800	889	\$1,654	897	\$1,377	1123	\$1,051	495	\$933	855	\$1,281	938

Table 2. Resource Consumption for IRL Upgrade

	1,2		2,3		2,4		2,5		4,5		5,6	
ITRL	Cost	Time	Cost	Time	Cost	Time	Cost	Time	Cost	Time	Cost	Time
1	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
2	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
3	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
4	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
5	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
6	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0	\$0	0
7	\$0	0	\$754	414	\$968	509	\$317	524	\$0		\$0	0
8	\$906	478	\$382	90	\$159	490	\$853	563	\$613	392	\$468	551
9	\$983	280	\$735	220	\$648	248	\$648	147	\$374	370	\$237	503
SSum	\$1,889	758	\$1,871	724	\$1,775	1247	\$1,818	1234	\$987	762	\$705	1054

With the proposed component Importance Measures for I^P , I^{CT} , and I^{LH} , this paper considers two scenarios for each measure to identify the importance of components towards achieving the KPP in question. While keeping all the other components constant, the two scenarios are to advance the current maturity of a component to (1) the next level, which is $\overline{TRL}_i = TRL_i + 1$ or $\overline{IRL}_{ij} = IRL_{ij} + 1$, and (2) increasing to its highest level, which is $\overline{TRL}_i = 9$ or $\overline{IRL}_{ij} = 9$.

Increasing Component Readiness by One Level

By increasing the component maturity by one level, Table 3 shows the results of the calculation. For the I^P component importance, Technology 2 is the most important component whose change in maturity has the largest impact on the maturity of the capability. When Technology 2 is increased by one level, the Capability SRL is upgraded from its current value of 0.622 to 0.646, and gives an I^P of 1.039. If the objective is to have the most increase in Capability SRL if only one component can be changed by one level, then Technology 2 is the most important one. The second and third most important components identified are Technologies 5 and 6, with an I^P of 1.031 and 1.021, respectively.

Table 3. Component Importance for the Scenario of Increasing by One Level

	Component	Current Readiness Level	SRL	I ^P	I ^P Rank	I ^{CT}	I ^{CT} Rank	I ^{LH}	I ^{LH} Rank
Technology	1	7	0.634	1.0195	5	1.2E-5	8	2.0E-5	8
	2	6	0.646	1.0390	1	4.2E-5	2	5.4E-5	2
	3	7	0.634	1.0189	6	1.2E-5	9	2.2E-5	6
	4	7	0.634	1.0197	4	2.7E-5	4	7.9E-5	1
	5	7	0.641	1.0307	2	4.3E-5	1	3.5E-5	3
	6	7	0.635	1.0207	3	3.1E-5	3	2.2E-5	4
Integration	1,2	7	0.631	1.0146	8	1.0E-5	11	1.9E-5	10
	2,3	6	0.631	1.0146	9	1.2E-5	10	2.2E-5	5
	2,4	6	0.629	1.0112	11	7.2E-6	12	1.4E-5	11
	2,5	6	0.628	1.0096	12	1.9E-5	6	1.1E-5	12
	4,5	7	0.631	1.0135	10	1.4E-5	7	2.1E-5	7
	5,6	7	0.633	1.0174	7	2.3E-5	5	2.0E-5	9

For the I^{CT} component importance, Technology 5 is the most importance with an I^{CT} of 4.3×10^{-5} indicating that the capability SRL will be increased by 4.3×10^{-5} for each dollar spent on maturing this technology. When considering budget allocation from a perspective of maturing the capability, Technology 5 is the most cost-effective component. The second and third most important components are Technologies 2 and 6.

Analyzing the I^{LH} component importance in the same way, we found that technology 4, with an I^{LH} of 7.9×10^{-5} has the most impact on capability. The capability SRL will be upgraded by 7.9×10^{-5} for every labor-hour spent on maturing this technology. When considering effort allocation from a perspective of maturing the capability, Technology 4 is the most effort-effective component. The second and third most important components are Technologies 2 and 5.

Figure 3 puts together the component importance evaluation from applying the three IMs to the capability of the system. The left vertical axis is the scale for I^P, and the right for I^{CT} and I^{LH}. Black bars represent the I^P importance with respect to the importance factor of TRL/IRL for the corresponding component, white bars for the I^{CT} importance with respect to cost, and grey bar for the I^{LH} importance with respect to effort. The higher the bar, the more important is that component with respect to the importance factor represented by the corresponding color.

Therefore, for the scenario of increasing by one level, Technologies 2, 5 and 6 are relatively more important than the other components with respect to TRL/IRL; Technologies 5, 2 and 6 are relatively more important than others with respect to cost; Technologies 4, 2 and 5 are relatively more important than others with respect to effort. When all three importance factors are considered simultaneously, Technologies 2, 4 and 5 are comparably more important components for the capability development within the system. Furthermore, Figure 3 implies, in general, that technologies are more important than integrations based on the current development maturity status of the system.

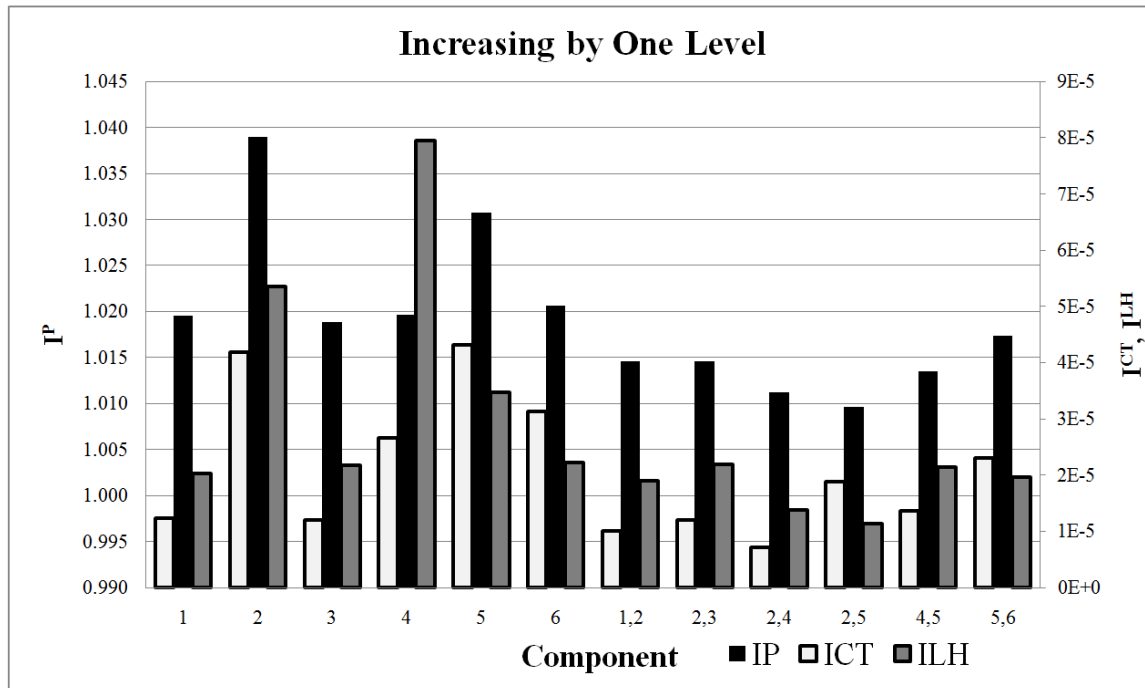


Figure 3. Component Importance Comparison for Increasing by One level Fully Maturing Components

For the scenario of increasing the component to its highest maturity level, Table 4 shows the results for considering each importance factor. Technology 2 is the most important component for all three factors, indicating the significant impact of fully maturing this technology on the maturity of the capability of the system. Therefore, resources must be prioritized towards the development of Technology 2 so as to ensure the satisfaction of the KPP of this system.

For the consideration of importance factor of TRL/IRL, Technology 5 and integration 2, 3 are the second and third most important components. Technology 5 and Integration 5, 6 are the second and third most important with respect to developmental cost. Technologies 4 and 5 are the second and third most important with respect to developmental effort. It should be noted here that some integrations also stand as very important components for maturing the capability to satisfy the KPP of the system.

Again, results of component importance calculation with respects to all three factors are plotted together in Figure 4 for comparison purposes.

Table 4. Component Importance for the Scenario of Increasing to the Most Maturity Level

	Component	Current Readiness Level	SRL	I ^P	I ^P Rank	I ^{CT}	I ^{CT} Rank	I ^{LH}	I ^{LH} Rank
Technology	1	7	0.646	1.0390	6	1.3E-5	9	2.7E-5	6
	2	6	0.695	1.1171	1	4.4E-5	1	8.1E-5	1
	3	7	0.646	1.0377	7	1.7E-5	6	2.1E-5	9
	4	7	0.647	1.0394	5	2.3E-5	4	4.9E-5	2
	5	7	0.660	1.0614	2	4.1E-5	2	4.5E-5	3
	6	7	0.648	1.0413	4	2.0E-5	5	2.7E-5	5
Integration	1,2	7	0.640	1.0291	10	9.6E-6	12	2.4E-5	7
	2,3	6	0.649	1.0437	3	1.5E-5	8	3.8E-5	4
	2,4	6	0.643	1.0337	9	1.2E-5	10	1.7E-5	11
	2,5	6	0.640	1.0288	11	9.8E-6	11	1.5E-5	12
	4,5	7	0.639	1.0270	12	1.7E-5	7	2.2E-5	8
	5,6	7	0.644	1.0347	8	3.1E-5	3	2.0E-5	10

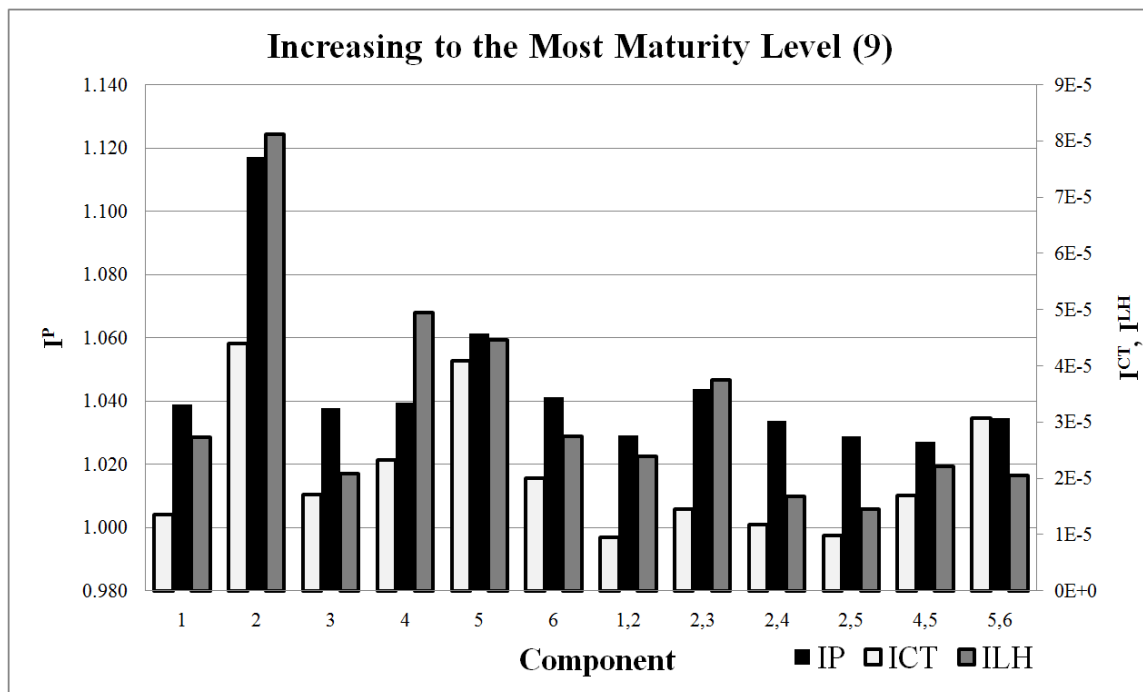


Figure 4. Component Importance Comparison for Increasing to the Most Maturity level

Conclusion

The complexity of developing systems that provide multiple functionalities and capabilities poses challenges to systems engineering managers. One of challenges is how to predict the development progress of the KPPs, and another one is how to leverage the allocation of resources to develop the Key Performance Parameter of interest. Volkert (2009) suggested a method for predicting the KPP development progress with the use of SRL. Based on his method, this paper proposes an approach for performing component importance analysis to identify the contributions of maturing components towards the maturity of a capability. Using their contributions as a guide, components can be ranked. The ranking can then serve as a guide for allocating resources when they are constrained. With the component importance quantified and identified, managers can make use of the information to prioritize available resources to the more important components, and thus to satisfy the preset development expectations.

Since TRL/IRL, developmental cost and effort are the major factors for maturing a system, and this paper proposes three corresponding importance measures (IMs). The application of these IMs to a notional example shows that the components' importance can be identified and distinguished. It was found that for this particular example, technology components are generally more important than integrations. This may be a reflection of the fact that the development of technologies usually starts first and integrations are considered later. However, a lot of systems cannot wait for integration until all technologies are completely matured. Therefore, even though development of integration may lag behind the development of technology, it is necessary to develop them in a parallel way. How the requirement of parallel development and implication from component importance analysis can jointly establish developmental strategy for determining resource allocation poses a question for future research.

Another fact to be noted from the definition and application of these IMs is that the component importance is identified based on the current development maturity status. What will importance rank change if the components that were identified to be very important have been matured? Will a spiral methodology be needed to address component importance for the long system development life cycle? Future research is needed to investigate these problems.

References

- Defense Acquisition University (DAU). (2009). *Manual for the operation of the joint capabilities integration and development system (JCIDS)*. Defense Acquisition University.
- Forbes, E., Volkert, R., et al. (2009). Implementation of a methodology supporting a comprehensive system-of-systems maturity analysis for use by the Littoral Combat Ship Mission Module program. In *Proceedings of the Sixth Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School.
- Magnaye, R., Sauser, B., et al. (2009). Using a system maturity scale to monitor and evaluate the development of systems. In *Proceedings of the Sixth Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School.
- Magnaye, R., Sauser, B., et al. (2010). Systems development planning using readiness levels in a cost of development minimization model. *Systems Engineering*, 13(4).



- Sauser, B., Magnaye, R., et al. (2010). Optimization of system maturity and equivalent system mass for exploration systems development. *2010 Conference on Systems Engineering Research (CSER)*, Hoboken, NJ.
- Sauser, B., Ramirez-Marquez, J., et al. (2008). A systems approach to expanding the technology readiness level within defense acquisition. *International Journal of Defense Acquisition Management*, 1, 39-58.
- Volkert, R. (2009, September 10). Notional assessment methodology for KPP accomplishment in a SoS proposed methodology for measuring performance progress within a system of systems (SoS). *SSC-Pacific*, Version 1.4.
- Zhang, P., Huo, C., et al. (2007). A novel measure of component importance considering cost for all-digital protection systems. Power Engineering Society General Meeting, IEEE.



Panel #23 – Using Knowledge Value Analysis and Real Options Analysis to Improve DoD Acquisition Planning and Management

Thursday, May 13, 2010

**3:30 p.m. –
5:00 p.m.**

Chair: *Jimmy D. Smith, Director, Advanced Water Sensors, PEO Integrated Warfare Systems*

Discussant: *Michael Schwind, Vice President, Maritime Solutions Sector, Siemens PLM*

Integrating System Dynamics Modeling and Knowledge Value Added for Improved Analysis of Alternatives: A Proof of Concept Study

David Ford, Texas A&M University, Tom Housel and John Dillard, Naval Postgraduate School

PEO-IWS ACB Insertion Portfolio Optimization

Johnathan Mun and Tom Housel, Naval Postgraduate School and Mark Wessman, Wessman Consultancy Group, Inc.



Integrating System Dynamics Modeling and Knowledge Value Added for Improved Analysis of Alternatives: A Proof of Concept Study

David Ford—Dr. David N. Ford is a professor in the Construction Engineering and Management Program, Zachry Department of Civil Engineering, Texas A&M University, and a Research Associate Professor of Acquisition at the Graduate School of Business and Public Policy at the US Naval Postgraduate School. He received BS and MS degrees from Tulane University and his PhD from MIT. Previously, he was on the faculty of the Department of Information Science, University of Bergen, Norway. For over 14 years, he designed and managed the development of facilities in industry and government. His current research interests include the project dynamics, managerial real options, and risk management.

Tom Housel—Dr. Tom Housel specializes in valuing intellectual capital, knowledge management, telecommunications, information technology, value-based business process reengineering, and knowledge value measurement in profit and non-profit organizations. He is a Full Professor for the Information Sciences (Systems) Department. His current research focuses on the use of KVA and “Real Options” models in identifying, valuing, maintaining, and exercising options in military decision-making. Prior to joining NPS, he was a Research Fellow for the Center for Telecommunications Management and Associate Professor at the Marshall School of Business at the USC. He received his PhD from the University of Utah in 1980.

John T. Dillard—COL John T. Dillard, US Army (Ret.), managed major weapons and communications programs for most of his 26-year career in the military, including development of the Javelin and Army Tactical Missile System (ATACMS) missile systems. Dillard now serves as a senior lecturer with the Graduate School of Business and Public Policy at the US Naval Postgraduate School in Monterey, CA. He is also a 1988 graduate of the Defense Systems Management College, Program Managers course. His current research interests include defense acquisition policy, organizational design and control, and product development. He is a member of PMI and other professional organizations.

Abstract

Effective and efficient DoD acquisition programs require the analysis of a wide range of materiel alternatives. Alternative diversity, difficulties in selecting metrics and measuring performance, and other factors make the Analysis of Alternatives (AoA) difficult. The benefits of alternatives should be included in AoA, but cost estimates predominate most AoA processes. Incorporating benefits into AoA is particularly difficult because of the intangible nature of many important benefits. The current work addresses the need to improve the use of benefits in AoA by building a system dynamics model of a military operation and integrating it with the Knowledge Value Added (KVA) methodology. The synergies may be able to significantly improve the accuracy of KVA estimates in the AoA process. A notional mobile weapon system was modeled and calibrated to reflect four weaponized Unmanned Aerial Vehicles (UAV). Modeling a hypothetical AoA for upgrading one of the UAV indicated that there were potentially significant synergies that can increase the number of alternatives that could be analyzed, establishing common units of benefit estimates for an AoA, improved reliability of an AoA, and improved justification of AoA results. These can improve alternative selection, thereby improving final materiel effectiveness, thereby improving DoD acquisition processes.



Introduction

The US defense acquisition process is initiated by the Joint Capabilities Integration and Development System (JCIDS), which is one of the three major decision support systems used in the DoD to interconnect and arrive at new warfighting capability. The JCIDS formulates force requirements with a “top down” approach that serves as both a Joint Force integrative process and one that can also hierarchically decompose the complexities of the battle spaces and their critical mission elements. It must also be aligned with the Planning, Programming, Budgeting and Execution System (PPBES) funding process as a way to descend from the strategic to the tactical in acquisition programs and budgets (*CJCSI 3170.01G*, 2009). JCIDS uses Capability Based Assessments (CBA) to validate capability gaps, to discover solutions such as those addressed by non-materiel-type changes to doctrine, organization, training, leadership & education, personnel, or facilities, or to pursue materiel solutions. Essential to CBA subprocesses is the knowledge of various functional warfighting Joint Capability Areas and how those communities operate within a joint paradigm. Functional Capability Boards are organized around top-tier functional areas such as Force Application, Logistics, etc.

Once needs are specifically derived in an area, and it is ascertained that they can only be addressed by new materiel, the acquisition community is still often left with either a variety of system types or technical approaches within a particular system type to fully address that capability need. An in-depth Analysis of Alternatives (AoA) helps sponsors and program managers compare options. Examples include manned or unmanned aircraft versus a missile, chemical energy versus kinetic energy kill mechanisms, etc. In the past, these have also been called Cost and Operational Effectiveness Analyses, and cost-effectiveness analyses—all variants of a Business Case Analysis.

The focus of the current work is on improving the Analysis of Alternatives (AoA) process that is used to make these major acquisition decisions. We will demonstrate the use of a system dynamics modeling approach that incorporates common units of benefits parameters using the KVA methodology and the potential improvements that might result in an AoA.

Problem Description

In typical weapon system acquisition programs, there is a point where an AoA is conducted to select the most viable and cost-effective materiel solution, so that it may be pursued into advanced development and production. Often, a selection for advanced technical development among competitive system prototypes is needed—a practice that has recently become official DoD acquisition policy, but is not a new idea (USD(AT&L) memo, September 19, 2007, Subj: Prototyping and Competition).

System concepts are further clarified during the Materiel Solutions Analysis phase and the accompanying AoA process. As part of these processes, programs move toward several kinds of evaluative cost comparisons that formulate costs estimates across a notional life cycle. In these early stages, programs use analogous and parametric cost estimation techniques. Parameters of system performance and key system characteristics are selected from technical and operational inputs. Usually, several Key Performance Parameters included in the Initial Capability Document are used in the AoA to quantify points of differentiation.

Simulations help address uncertainty across likely operational contingencies (Ford & Dillard, 2008; 2009, May 13-14; 2009, July). Balancing of programmatic and operational



risks should be accounted for, but costs predominate most analyses because, for major weapon systems, they can be huge. Not only must research and development costs be considered, but also the production costs, and, beyond that, all of the operating costs of spares, diagnostics, maintenance, tools, training manuals, etc., must be estimated.

While the emphasis is clearly on cost in these stages, operational effectiveness must also be considered because that is where benefits are realized. Current guidance does not provide a method for estimating benefits in common units. Some feel that the emphasis is disproportionately on cost without enough emphasis on benefits. In current guidance (e.g., the GAO's Cost Estimation Guide, 2009, p. 35) benefits are primarily in the form of cost avoidance or cost savings. Clearly, these are not equivalent to normal estimates of benefits in the business world where revenue is the primary indicator of benefit and is not derived from the denominator or cost side of the equation. Monetizing benefits as some form of cost savings/avoidance leads to a slippery slope where the only indicator of value, or numerator in a productivity equation, such as return on investment (ROI), is a derivative of the denominator, i.e., cost. Such predilections inevitably lead to the lowest cost alternatives, which may not provide the highest benefits.

Research Focus

Benefits, in common units, should be included in AoA to enable higher fidelity comparisons among alternatives on the basis of value and not just cost. But how can sponsors and program managers best value very real and important but intangible benefits such as combat effectiveness, survivability, or national security? Lacking a credible ability to quantify such subjective or intangible benefits of the capabilities of a system type (or technical alternative) is a serious omission in any rigorous analysis of alternatives. The third author's experience includes several recent examples that illustrate the need for more than a conventional cost effectiveness analysis to defend a program requirement, or a system parameter of technical capability. Often, a particular system parameter of capability (e.g., weight, C-130 transportability, vertical take-off and landing) become a metric of program life and death, but with notably sparse articulation of empirical benefit to the customer/end user.

The Case of the Javelin Anti-Tank Weapon System

The Javelin anti-tank weapon system was, when it was conceptualized, merely named after its requirement as the Advanced Anti-Armor Weapon System–Medium (AAWS–M). In 1987-89, the US Army tested three competing technologies to fulfill the operational need for a one-man-portable anti-armor weapon system in the medium range (1,000–2,000 meter) category, and replace its aging and ineffective DRAGON weapon system. Principally, the weapon was to do the following: be able to defeat current and projected threat armored vehicles (including tanks), have a maximum range of at least 2,000 meters, weigh no more than 20.5 kg (with under 15.5 kg being desirable), have the ability to be fired from enclosed spaces, and be able to engage armored vehicles under cover or in hull defilade. The US Marine Corps agreed to these requirements, promising to pay for production items, but not to fund research and development.

In August of 1986, "Proof of Principle" contracts of \$30 million each were awarded to three competing contractor teams, spanning a 27-month period (the phase we now call "Technology Development") to develop the technologies and conduct a "fly-off" missile competition. Each offered the needed capability solutions with differing technologies. Ford Aerospace teamed with its partner Loral Systems, offering a laser beam-riding missile.



Hughes Aircraft teamed with Boeing to offer a fiber-optic guided missile. Texas Instruments teamed with Martin-Marietta, offering an imaging infra-red (I2R) or forward looking infra-red (FLIR) missile system. Each candidate system also offered some specific operational advantages and disadvantages that were almost impossible to quantify in terms of cost:

- The Ford/Loral Laser Beam Rider required an exposed gunner and man-in-loop throughout its rapid flight. It was cheapest at an estimated \$90,000 “cost per kill,” a figure comprised not only of average production unit cost estimates but also reliability and accuracy estimates. It was fairly effective in terms of potential combat utility, with diminishing probability-of-hit at increasing range.
- The Hughes/Boeing Fiber-Optic guided prototype enabled an unexposed gunner (once launched) and also required man-in-loop throughout its slower flight. It was costlier, but less affected by range accuracy with its automatic lock-on and guidance in its terminal stage of flight, and even offered target switching. It was also more gunner training (learning) intensive, but could attack targets from above where their armor was thinnest.
- The FLIR prototype offered completely autonomous “fire and forget” flight to target after launch, was perceived as both costliest and technologically riskiest. It would be easiest to train and would be effective to maximum ranges by means of its target acquisition sensor and guidance packages. It was an outgrowth of a 1980 initiative by the Defense Advanced Research Project Agency (DARPA) called Tank Breaker that also used “top attack” as a more effective means of armored target defeat.

1988 was a busy year for the AAWS-M industry contractors, as well as the government acquirers and program sponsors. All three candidate teams engaged with finally building and flight-testing their missile prototypes. They were also submitting their bids to the government’s Request for Proposal for the upcoming advanced development phase. On the government side, acquirers were evaluating these bids and preparing to award the 36-month Engineering and Manufacturing Development (EMD) phase contract, while sponsors were completing a Cost and Operational Effectiveness Analysis (COEA) of the three candidate AAWS-M materiel solutions. Each of the teams enjoyed generally successful missile flight test outcomes as the “proof of principle” phase ended. Each flew over a dozen missiles and achieved a target hit rate of over 60%.

The Laser Beam Rider candidate emerged the winner of the COEA, presumably from weighted cost/efficiency factors. But in a strange twist, the concurrent deliberation of the Source Selection Evaluation Board (SSEB) instead chose the FLIR candidate, presumably because of a bias toward “fire & forget.” As part of a typical capability formulation process, technical constraints are deliberately avoided in requirements documents, to allow and encourage a maximum range of alternative solutions to the need or capability deficiency. While time of flight and gunner survivability were not stated requirements in the AAWS-M Joint Required Operational Capability document per se, “fire and forget” nevertheless translated into greatly enhanced gunner survivability, and overwhelmingly appealed to user representatives (and government developers).

The EMD contract was awarded in June of 1989 to the Joint Venture team of Texas Instruments and Martin-Marietta. However, at about 18 months into this program, serious technical problems doubled the expected cost of development and added about eighteen more months to the originally planned thirty-six months to complete. This constituted a Nunn-McCurdy breach of cost and schedule thresholds, with requisite Congressional notifications and formal re-baselining taking the better part of the next year to accomplish.



Various technical issues plagued the program at this point, with system weight being perhaps chief among them. User representatives convened a Joint Requirements Overview Council (JROC) to re-evaluate the maximum weight requirement of 45 pounds and increased the program threshold to 49.5 pounds. Clearly, the Army and Marine Corps communities wanted the emergent system and its planned capabilities. But that didn't resolve all of AAWS-M's issues.

During the months that the program teetered on the brink of termination for its technical and business issues, the Director of the OSD Office of Program Analysis and Evaluation (DPAE), as a principle member of the DAB, took the program to task stating that if the FLIR version could not be shown able to achieve the same \$90K "cost per kill" as had been estimated for the laser beam-rider, then the program should be terminated and re-started, changing technologies and pursuing the less risky laser-guided version. The principal cost driver of the FLIR technology that enabled fire-and-forget was a 64x64 matrix (of heat detectors/pixels) focal plane array (FPA), to be manufactured by one of the Joint Venture partners. These tiny micro-chips would comprise almost 14% of the estimated average unit production cost (UPC) of the entire missile. The ability of one of the few producers in the world to produce them with economically sufficient yield, and to achieve their rigorous performance specifications for sensitivity seemed, for a while, to hold the fate of the entire program. Intense scrutiny of projected yields and production costs of these critical components would determine whether the program was feasible from this aspect alone, some believed. But the answer was somewhat ambiguous, with roughly \$12k being the target for average UPC, given a planned buy quantity of about 70,000. And cost of FPAs wasn't the only problem with them. But it turned out that their benefits could be described in a fairly tangible way.

The AAWS-M FPA specifications were derived from a scenario-based target list of potential threat vehicles in different environments of atmospheric temperature, humidity, obscuration, etc. When the user community saw that early developmental AAWS-M focal plane arrays were not meeting the full specifications, they convened another JROC to allow stepped, incremental achievement of target defeat scenarios over time—something we would now refer to as evolutionary growth. They stratified performance in terms of levels A, B and C to convey degrees of target defeat capability in FPAs—a very unusual move by sponsors, having to dissect a requirement to accommodate the pace of technological achievement.⁶⁵ This provided a qualitative assessment of what was achievable and satisfactory for system performance. Once again, the communities that needed AAWS-M's capabilities were trying to ease the path forward.

Fortunately, independent program evaluation teams also reported that FLIR technology was progressing and would be achievable within a re-baselined program. This joint position, along with wider program advocacy, curtailed the technical and business arguments and the fire-and-forget Javelin was allowed to proceed. An additional and more capable provider of FPAs was brought in and accelerated as a second source for this critical component. After still more and difficult advanced development program challenges, AAWS-M eventually became the Javelin—and is known today as one of our most successful combat systems. (In the end, soldiers and Marines never had to accept to accept B and C-level FPA performance, as the full-capable FPA technology did in fact emerge in time for

⁶⁵ Perhaps not unlike today's emergence of an Apple iPhone® being followed soon after by release of a 3G-capable iPhone®.



fielding. And system weight has been held just below 49.5 pounds throughout its many years of production.)

There are many business and public policy lessons to be learned from the Javelin program. Within its long saga from initial concept to modern-day deployment and combat use are illustrations of requirements capture, early prototyping, technology readiness, modeling and simulation, economic forces of competition, acquisition strategy, decision bureaucracy, product discovery, and economies of scale, etc. Perhaps the best lesson learned from the case presented here about analyzing alternatives is that a single, unstated, qualitative factor of performance (gunner survivability) ultimately drove the choice. Javelin had a requirements document with many pages of quantifiable requirements stated as measures of performance and effectiveness. But the parameter of system technology that promised the most of what was impossible to quantify became the overriding factor in the selection of alternatives. A magazine advertisement purchased by the Joint Venture shortly after their EMD contract win said it eloquently: “Fire & Forget AAWS-M: The Gunner Wins.” The failure of the Javelin program to move to the final solution faster and more directly is due in large part to the insufficient articulation of benefits as part of the Analysis of Alternatives process.

Research Question

As illustrated by the Javelin program, there is a basic need for the use of a common units of benefit estimate in the Analysis of Alternatives process. This should lead to including common units benefits estimates as well as costs in the acquisition AoAs. The problem is to develop a means to do this more effectively, given the nebulous nature of so many of the critical benefits of weapon systems. How can such a method be consistently applied to many alternatives across a wide range of operational conditions? ***The current research examines how KVA can be integrated with system dynamics modeling to generate defensible common units of benefit estimates that will improve the rigor of the AoA process and thereby improve acquisition processes.***

The goals of the current work are:

- Examine how military operations systems dynamics simulations can be combined with the KVA approach,
- Identify potential advantages and disadvantages of integrating military operations simulations and the KVA approach,
- Investigate the potential of exploiting the benefits from the synergy of SD and KVA to improve acquisition AoA processes, and
- Identify and describe potential implications of the integration on acquisition practice.

Due to the preliminary nature of this proof-of-concept study, precise descriptions of system operations are necessary. The focus is on the potential usefulness of integrating SD and KVA.

Introduction to Knowledge Value Analysis

In the US Military context, the Knowledge Value Added (KVA) methodology is a new way of approaching the problems of estimating the productivity (e.g., in terms of ROI) for military capabilities embedded in processes such as the CONOPS for a weapons system. In the current study, we posited several alternative CONOPS for a UAV system and used



system dynamic modeling to evaluate their relative productivity. The KVA approach was used to estimate the parameters based on the system dynamic models by providing the estimates of the relative productivity (i.e., the ROI⁶⁶) of each alternative.

In a broader context, KVA also addresses the requirements of the many Department of Defense (DoD) policies and directives previously reviewed by providing a means to generate comparable value or benefit estimates for various processes and the technologies and people that execute them. It does this by providing a common and relatively objective means to estimate the value of new technologies as required in the:

- *Clinger-Cohen Act of 1996* that mandates the assessment of the cost benefits for information technology investments.
- Government Accountability Office's *Assessing Risks and Returns: A Guide for Evaluating Federal Agencies' IT Investment Decision-Making*, Version 1, (February 1997) that requires that IT investments apply ROI measures.
- *DoD Directive 8115.01*, issued October 2005, that mandates the use of performance metrics based on outputs, with ROI analysis required for all current and planned IT investments.
- *DoD Risk Management Guidance Defense Acquisition* guide book that requires alternatives to the traditional cost estimation be considered because legacy cost models tend not to adequately address costs associated with information systems or the risks associated with them.

KVA is a methodology that describes all organizational outputs in common units. This provides a means to compare the outputs of all assets (human, machine, information technology) regardless of the aggregated outputs produced. Thus, it provides insights about the productivity level of processes, people, and systems in terms of a ratio of common units of output produced by each asset (a measure of benefits) divided by the cost to produce the output. By capturing the value of knowledge embedded in an organization's core processes, employees and technology, KVA identifies the actual cost and value of a people, systems, or processes. Because KVA identifies every process required to produce an output and the historical costs of those processes, unit costs and unit values of outputs, processes, functions or services are calculated. An output is defined as the end-result of an organization's operations; it can be a product or service, as shown in Figure 1.

⁶⁶ ROI is defined as the revenue-cost/cost where revenue is defined as the price per common unit of benefit using a market comparables approach. Given that the price per common unit is a constant, precision in estimating the market comparable price, i.e., revenue, is not required.



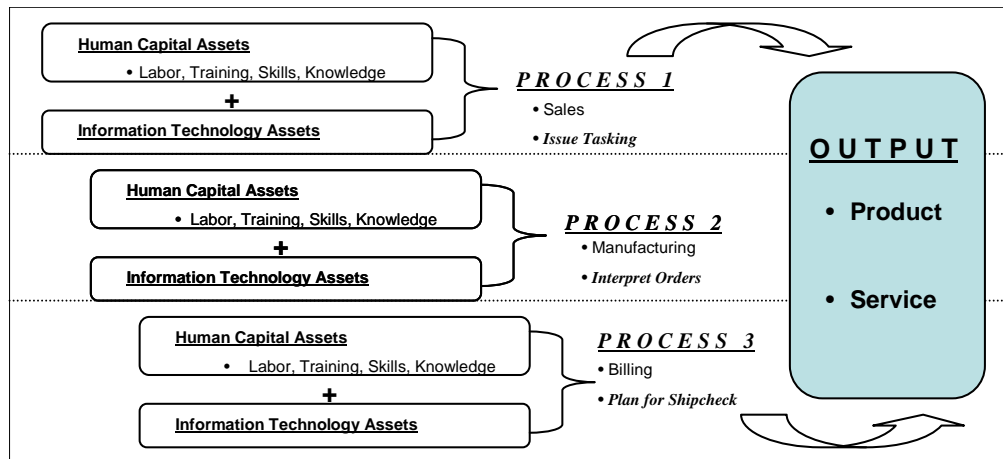


Figure 1. Measuring Output

For the purpose of the systems dynamics model developed for this study, KVA was used to describe the outputs of all the processes and subprocesses in common units. This allowed us to make their relative performance (e.g., productivity, ROIs) comparable. KVA was used to measure the value added by the human capital assets (i.e., military personnel executing the processes) and the system assets by analyzing the processes performances. KVA provided a means to set the systems dynamic model parameters so that the results would provide a means to compare the performance of various approaches to the system problem.

By capturing the value of knowledge embedded in systems and in use in operators of the processes, KVA identified the productivity of the system-process alternatives. Because KVA identified every process output required to produce the final aggregated output, the common unit costs and the common unit values were estimated. This allowed for the benchmarking of various systems and the processes they support with any other similar processes across the military.

The KVA methodology has been applied in over 80 projects within the DoD, from flight scheduling applications to ship maintenance and modernization processes to the current project analyzing several alternative approaches to the system alternatives problem. In general, the KVA methodology was used for this study because it could:

- Compare alternative approaches modeled with a systems dynamics model in terms of their relative productivity,
- Allocate value and costs to common units of output,
- Measure value added by the system alternatives based on the outputs each produced, and
- Relate outputs to cost of producing those outputs in common units.

Describing processes in common units also permits, but does not require, market comparable data to be generated, particularly important for non-profits like the US Military. Using a market comparables approach, data from the commercial sector can be used to estimate price per common unit, allowing for revenue estimates of process outputs for non-profits. This also provides a common units basis to define benefit streams, regardless of the process analyzed.

KVA differs from other nonprofit ROI models because it can allow for revenue estimates, enabling the use of traditional accounting, financial performance, and profitability measures at the sub-organizational level. KVA can rank processes or process alternatives by their relative ROIs. This assists decision-makers in identifying how much various processes or process alternatives add value.

In KVA, value is quantified in two key metrics: Return-on-Knowledge (ROK: revenue/cost) and ROI (revenue-investment cost/investment cost). The raw data from a KVA analysis can become the input into the ROI models and various forecasting techniques such as real options analysis, portfolio optimization, Monte Carlo simulation. By tracking the historical volatility of price and cost per unit as well as ROI, it is possible to establish risk (as compared to uncertainty) distributions, which is important for accurately estimating the forecasted values for portfolio optimization and real options analysis.

Introduction to System Dynamics

The system dynamics methodology applies a control theory perspective to the design and management of complex human systems. System dynamics combines servo-mechanism thinking with computer simulation to analyze systems. It is one of several established and successful approaches to systems analysis and design (Flood & Jackson, 1991; Lane & Jackson, 1995; Jackson, 2003). Forrester (1961) develops the methodology's philosophy, and Sterman (2000) specifies the modeling process with examples and describes numerous applications. The methodology has been extensively used for this purpose, including studying development projects. The system dynamics perspective focuses on how the internal structure of a system impacts system and managerial behavior and, thereby, performance over time. The approach is unique in its integrated use of stocks and flows, causal feedback, and time delays to model and explain processes, resources, information, and management policies. Stocks represent accumulations or backlogs of work, people, information, or other portions of the system that change over time. Flows represent the movement of those commodities into, between, and out of stocks. The methodology's ability to model many diverse system components (e.g., work, people, money, value), processes (e.g., design, technology development, production, operations, quality assurance), and managerial decision-making and actions (e.g., forecasting, resource allocation) makes system dynamics useful for modeling and investigating military operations, the design of materiel, and acquisition.

When applied to acquisition programs, system dynamics has focused on how performance evolves in response to interactions among development strategy (e.g., evolutionary development versus traditional), managerial decision-making (e.g., scope developed in specific blocks), and development processes (e.g., concurrence). System dynamics is appropriate for modeling acquisition because of its ability to explicitly model critical aspects of development projects. System dynamics models of development projects are purposefully simple relative to actual practice to expose the relationships between causal structures and the behavior and performance that they create. Therefore, although many processes and features of system design and participants interact to determine performance, only those that describe features related to the topic of study are included. The importance of deleted features can be tested when system dynamics is used to test the ability of the model structure to explain system behavior and performance.

Based on the preceding and the authors' experience with system dynamics, there appears to be an opportunity to exploit the capabilities of the system dynamics methodology to make the Knowledge Value Added approach more accurate.



Research Methodology

In the current work, Knowledge Value Added and system dynamics were integrated to test their ability to improve the precision of AoAs in acquisition programs. A generic structure of a mobile weapon system process was first developed and tested using the system dynamics methodology. Then, KVA value and cost estimates were operationalized in the system dynamics model. The model was calibrated to reflect four extant weaponized Unmanned Aerial Vehicles (UAVs). One of those calibrations was used as the basis for using the model in a hypothetical AoA for upgrading the UAV to address a different type of target. Simulation results were analyzed to test the ability of the system dynamics model to estimate benefits streams using KVA in terms of the relative value added of the capabilities of the system.

A Generic Model of Mobile Weapons Use

The model has three sectors: weapons movement, target evolution, and KVA analysis. As will be described, the model structure simulates two critical aspects of mobile weapon system operations: 1) the support and movement of the weapon and 2) target evolution from identification through confirmation of destruction.

The Weapons Movement Sector

The Weapons Movement sector of the model simulates the positions and movements of weapons (e.g., individual UAVs or Javelin gunners). Figure 2 shows the positions that weapons (generically called “assets”) can take (boxes) and the rates of their movements from one position to another (arrows between boxes). It is assumed that the total number of assets remains constant, i.e., no weapons are added or lost during operations. This assumption can be relaxed when modeling a specific asset. The movement of weapons is a subprocess of operating the weapon system that adds value and imbeds learning into tools, requires learning time for operators to be capable of doing, and requires processing time to accomplish. Therefore, the completion of moving weapons to the station and back to the base is an output of that subprocess and an input to the KVA analysis. The combination of the two movements “Assets arrive at station rate” and “Assets arriving as base rate” represent the accomplishment of the vehicle movement subprocess.

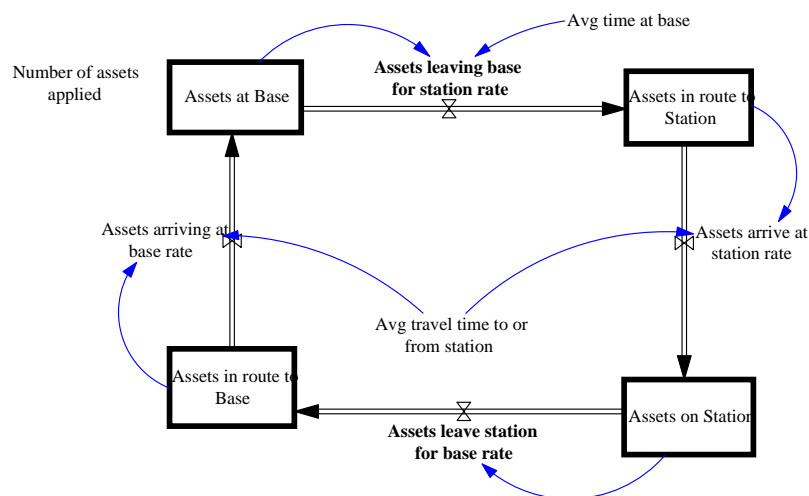


Figure 2. Positions and Movement of Weapons during Operations

Each rate in the weapons sector describes the average movement of the weapons in the accumulation that precedes the rate. Each rate is defined with the number of weapons preparing for that rate (to leave base or station) or event (to arrive at station or base) and the average time spent by a weapon in the preceding accumulation. For example, the (average) “Assets leaving base for station” rate is equal to the number of assets at the base divided by the average time that a weapon spends at the base between trips to the station. This formulation increases the average departure rate with more weapons at the base and decreases the average departure rate if weapons stay at the base longer. The average time at the base is characteristic of particular assets and can generate different behaviors and performances across weapons and configurations.

The Target Evolution Sector

The target evolution sector of the model simulates the development of targets through five subprocesses of system operations.

1. **Acquire target:** Includes detection, recognition, location, classification (identification), and confirmation (Lombardo, 2003; Global Security, 2010).
2. **Fire support coordination:** Allocates targets to weapons by a group of people that have access to information about the battlefield situation, and doctrine, major systems, significant capabilities and limitations and often their TTP [tactics, techniques and procedures] (Williams, 2001).
3. **Fire mission development:** Prepares specific instructions and target information for transmission to the weapons team and to the weapon (e.g., target location coordinates).
4. **Engage target:** Weapons operators (e.g., pilots for UAV) maneuver the weapon within striking distance of the target, enter the target coordinates and launch munitions .
5. **Battlefield assessment:** Often the same asset as was used for target acquisition is used to evaluate the success of engagement in destroying the target.

In the model, targets evolve through these stages in an “aging chain” structure of sequential accumulations (backlogs + work in progress, referred to here as backlogs) and (sub)processes that drain those backlogs and contribute to the backlog of the next downstream subprocess. Figure 3 shows the conditions of targets (boxes) and the rates of their movements from one condition to another (arrows between boxes) due to subprocesses. The movements “Acquire target completion rate,” “Fire support coordination to asset,” “Fire mission completion rate,” “Engage target,” and “Battlefield assessment rate” are subprocesses that add value, imbed learning in tools, require learning time for operators to be capable of doing, and require processing time to complete. Therefore, they are each outputs of those subprocesses and inputs to the KVA analysis.

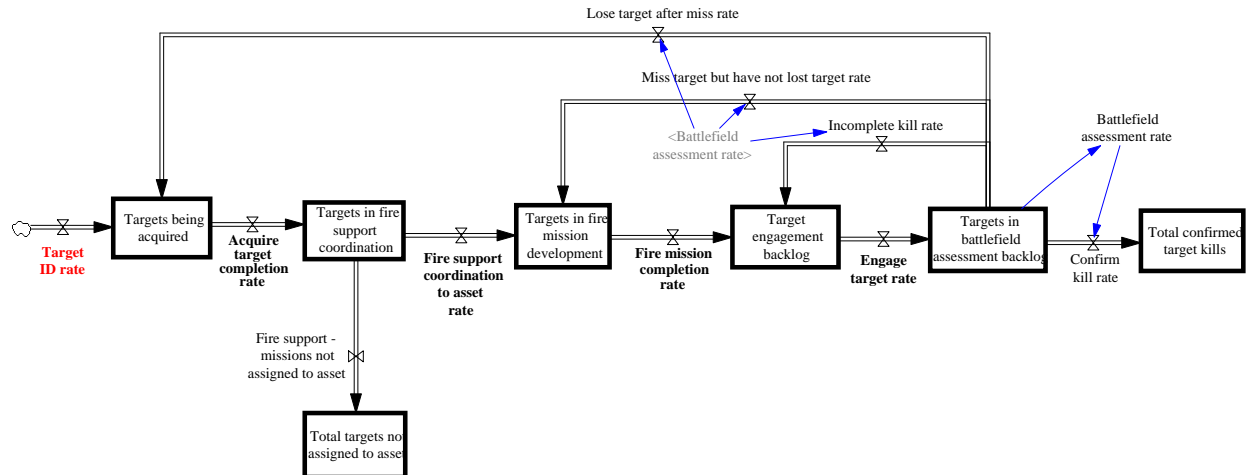


Figure 3. Accumulations and Movements of Targets in Weapon System Operations

In addition to the primary flows of targets through the subprocesses, the targets sector models three common causes of mission failure: 1) hitting the target but failing to destroy it, 2) missing the target, and 3) missing the target and losing the location information needed to engage the target again (e.g., because the target moved). Each cause moves the target to a different condition in the target aging chain. Hitting the target but failing to destroy it (e.g., a hardened target) requires reengagement but often no additional targeting information. After Battlefield assessment, these targets are returned to the Target engagement backlog. Missing the target (e.g., a small target) requires that the fire mission be developed again to re-aim the weapon prior to reengagement. Therefore, these targets are returned to the Targets in fire mission development backlog after Battlefield assessment. Losing the target (e.g., a fast moving vehicle) requires that the target be reacquired. Therefore, these targets are returned to the Targets being acquired backlog after Battlefield assessment.

In a manner similar to the modeling of the movement of weapons, the rates in the targets sector describe the average movement of targets between backlogs. The primary rates in the aging chain are defined by the number of targets in the backlogs of the subprocess and the average time required to perform the subprocess. The average time required to perform the subprocess is characteristic of particular subprocesses (e.g., different engagement durations for different weapons) and can generate different behavior and performance across weapons and configurations. When two or more flows drain a backlog (Targets in fire support coordination and Battlefield assessment backlog) the total outflow is split between the flows with a percent that leaves the stock through each outflow. The return flows are each a fraction of the Battlefield assessment rate. Those fractions are based on the ability of the weapon to successfully destroy, hit, and not lose targets. Therefore, like in practice, different weapon alternatives (e.g., range, payload, dash speed) impact mission success. The model use section of this report describes how these features of the model were used to describe operational scenarios and weapon configurations. The fraction of the targets that are returned due to being hit but not destroyed (to engagement backlog), missed (to mission development backlog), or lost (to target acquisition backlog) battlefield assessment rate are described with the probability of destruction if the target is hit with the ordnance ($p(\text{kill if hit})$) or $p(\text{kill})$, the probability of the weapon hitting the target with ordnance ($p(\text{hit})$), and the probability of not losing the target if it is missed with the

ordinance ($p(\text{not lose})$), respectively.⁶⁷ The probabilities are determined by comparing the ability of a weapon to successfully destroy, hit, and not lose targets to the characteristics of the target and a function that describes how the weapon's ability compared to the target impacts weapon performance. More specifically, the probability of kill is modeled with the weapon's payload compared to the lethal payload (i.e., ordinance size required to kill); probability of hit is modeled with the weapon's dash speed compared to the target's speed; and the probability of not losing the target is modeled with the weapon's range compared to the target's distance from the base. Therefore:

$$\begin{aligned} p(\text{kill}) &= f_k(\text{Payload} / \text{Lethal payload}) \\ p(\text{hit}) &= f_h(\text{Dash speed} / \text{Target speed}) \\ p(\text{not lose}) &= f_{nl}(\text{Range} / \text{Target distance from base}) \end{aligned}$$

where:

$p(\text{kill})$ - probability of destruction if the target is hit with the ordinance
 $p(\text{hit})$ - probability of the weapon hitting the target with ordinance
 $p(\text{not lose})$ - probability of not losing the target if it is missed with the ordinance

The three functions that estimate the probabilities based on the ratios are assumed to be simple but realistic relations that include the entire range of possible conditions.⁶⁸ The function relating the Payload/Lethal payload ratio to the probability of kill is assumed to increase linearly from $p(\text{kill})=0$ when the ratio is zero (i.e., no payload prevents any chance of target destruction) to $p(\text{kill})=100\%$ when the ratio is greater than or equal to 1 (i.e., if the payload exceeds the lethal payload the target is assumed to be destroyed if it is hit). The function relating the Dash speed/Target speed ratio to the probability of the weapon hitting the target with ordinance assumes that the vehicle will “chase” a moving target and that the faster the vehicle is then the closer it can get to the target before releasing ordinance, increasing the likelihood of hitting the target with the ordinance. However, there is always some possibility of missing a target even if the vehicle is faster than and, therefore, close to the target. The function is assumed to have an elongated “S” shape from $p(\text{hit})=0$ when the ratio is zero (i.e., no Dash speed prevents hitting the target) to $p(\text{hit})=90\%$ when the ratio is greater than or equal to three (i.e., high likelihood of hit if weapon speed far exceeds target speed).⁶⁹ The function relating the Range/Target distance from base ratio to the probability of not losing the target if it is missed with the ordinance assumes that the vehicle will move toward the target but that the target may also move, sometimes closer to the vehicle and sometimes away from it. When the target moves away from the vehicle it may move out of the vehicle's range, causing the vehicle to lose the target. The function is assumed to have a stretched out “S” shape from $p(\text{not lose})=0$ when the ratio is zero (i.e., no weapon range causes the vehicle to always lose the target) to $p(\text{not lose})=95\%$ when the ratio is greater than or equal to 1.8, reflecting some chance of losing the target even if it is well within the vehicle's range.

⁶⁷ The probability of not losing a target is used instead of the probability of losing a target to retain a “bigger is better” standard for all three measures and, therefore, facilitate intuitive understanding of the model.

⁶⁸ These functions can be described more accurately with additional weapon testing information.

⁶⁹ The assumed function relating the dash speed/target speed to the probability of hitting the target is probably lower than current experience, but is used to reflect the change in targets described in the Model Use section of this report.



The KVA Sector

The KVA metrics were fully operationalized within the system dynamics model. The KVA sector uses operations information from the weapons and targets sectors of the model and characteristic descriptors of weapons to generate relative value metrics for each subprocess (including weapons capability outputs) of the UAV operations. KVA generates a productivity ratio that reflects output/input. If monetized, this ratio can be a traditional benefit:cost ratio (e.g., ROI if benefits are monetized as a form of revenue surrogate). Other measures of benefits and costs can also be used, as long as there are common units in the numerator (benefits) and all the cost units of all the contributors to the denominator are the same (as is most often the case because costs are almost always monetized), so they can each be aggregated. At each point in time, each subprocess's productivity is the benefits it has generated divided by the costs to generate those benefits (i.e., output/input). The model includes both monetized and time-based KVA metrics. However, it was decided that the current model would be simpler to interpret by using the non-monetized common units of output (as described in terms of the units of time it would take the average person to learn how to produce the outputs) as the numerator. This results in the output (i.e., common units of learning time)/input (cost to produce the outputs) standard definition of productivity.

A time-based application of KVA uses Learning Time to quantify benefits and Touch Time to quantify costs. Learning time captures the benefits derived from human processing (e.g., flying the vehicle), automated processes (e.g., take offs and landings), and (importantly) technologies integrated into the weapon. One of several ways to quantify a subprocess's Learning Time is to estimate the average time required for a common point of reference learner to be trained and become competent in performing the subprocess. Each subprocess is assigned a Unit Learning Time that reflects the relative (compared to other subprocesses) complexity of the subprocess.

The denominator of KVA subprocess productivity ratios represents subprocess costs. Each subprocess is assigned a Unit Touch Time that reflects the relative (compared to other subprocesses) effort required to perform the subprocess. The total time spent performing a subprocess at any given time is the product of the average time required for the subprocess and the number of performances of the subprocess. Therefore, a set of equations for estimating the costs of a single subprocess in each time period are:

$$\text{Subprocess learning Time accumulated to date} = \sum (\text{Rate of subprocess operation} * \text{Subprocess unit learning time}) * dt$$

$$\text{Subprocess touch time accumulated to date} = \sum (\text{Rate of subprocess operation} * \text{Subprocess unit touch time}) * dt$$

$$\text{Subprocess Productivity} = \frac{\text{Subprocess learning time accumulated to date}}{\text{Subprocess touch time accumulated to date}}$$

Learning Times and Touch Times are also aggregated across subprocesses to estimate the productivity of the entire operation. This allows the comparison of different asset configurations (alternatives).

Model Calibration and Testing

The KVA+SD model was calibrated to the operations of four actual weaponized UAVs. The operations included the following subprocesses:

1. acquire target



2. fire support coordination
3. fire mission development
4. move weapons
5. engage targets
6. battlefield assessments

Three of the four actual UAVs are operational: Predator, Sky Warrior, and Reaper. The fourth UAV is the X-47B. Basic characteristics relating to performance were collected for the four UAVs from publicly available sources (e.g., Global Security, 2010). That information included vehicle range, total mission time, time on station, dash speed, and payload. In some cases, multiple versions of the vehicle with different characteristics have been developed. In these cases, a single version was selected and used. Other information was estimated for each vehicle's operations, including learning and processing times for each subprocess. Reasonable assumptions were used in making these estimates, such as that the time required to engage a target after arriving on station is inversely proportional to the vehicle's dash speed (i.e., faster dash speeds reduce the time required to engage). These estimates are rough, but adequate for this proof-of-concept study, which seeks to determine if the model is capable of reflecting differences in characteristics in KVA parameters, not predict actual outcomes.

The model was tested using standard tests for system dynamics models (Forrester & Senge, 1980; Sterman, 2000), including for structural similarity to the actual system, reasonable behavior over a wide range of input values, and behavior similarity to actual systems. Basing the model on previously validated models, the literature improves the model's structural similarity to actual acquisition projects, as practiced. Model behavior (e.g., simulated sizes of backlogs for subprocesses and rates of performing operations) were compared to typical behavior and found to be similar. For example, before operations start at the beginning of the operational scenario (described in what follows) the backlogs are empty and no operations are being performed. The appearance of targets increases subprocess backlogs and rates of operation as weapons leave base and subsequently arrive on station, targets are acquired, fire coordinated, missions developed, targets engaged, and the battlefield assessed.

In the evolution of targets, these backlogs and subprocesses increase sequentially through the series of operations. The growth of operations and backlogs slows as capacities adjust to demand (backlog sizes), until the operations are in dynamic equilibrium conditions with sizes of backlogs and operations rates remaining within a relatively narrow range. This represents "steady state" operations that could be continued for a significant period of time, e.g., until damage to weapons or maintenance (not included in the current model) change weapon availability. Model behavior was also tested with extreme input values such as perfect operations (e.g., probability of hit=100%) and very large versus very small number of weapons and targets as well as more typical conditions. Model behavior remained defensible across wide ranges of input values, including extreme values. These tests increase confidence that the model generates realistic operational process behavior patterns due to the same causal relations found in the type of operations investigated (i.e., generates "the right behavior for the right reasons").



The operational scenario was described with the quantity and characteristics of the targets.⁷⁰ A stream of targets entered the target acquisition backlog at a steady rate of five targets per minute. The target distance from the base was assumed to vary uniformly from 400-1100nm. This describes targets that range from being closer to the weapon's base than the shortest weapon's range to targets that are farther from the base than the longest weapon's range. The speed of the targets was assumed to vary uniformly from 50 to 250 nm. This describes targets that range from those that are immobile to targets that are faster than the fastest weapon. The payload required to destroy the target if hit (i.e., lethal payload) was assumed to vary uniformly from 400 to 1,000 lbs. This describes targets that range from being very soft to very hardened.

KVA productivities for the six subprocesses and the cumulative for those processes for the four Weaponized UAVs are shown in Table 1. For example, the KVA productivity ratios for the Fire Mission Development subprocess for the four UAV are 943 (Predator), 3,122 (Reaper), 1,222 (Sky Warrior), and 3,962 (X-47B). They represent the benefits (output) per unit of cost (input) and, therefore, can also be interpreted as a measure of the return on the investment, in percent. These values remain constant in the model after steady state operations have been established. As an example of the components of the ratios, the Fire Mission Development subprocess ratio for the Predator (943) is the quotient of the accumulated benefits (e.g., after 5 hours of operations) of 79,684 learning-time hours and 84.5 processing-time hours. In the simulated steady state operations this accumulated learning time hours increases at a rate of 301 learning-time hours per minute (the product of the estimated 500 learning-time hours per fire development operation and an average fire development rate of 0.6 targets developed per minute) and the processing-time hours increases at a rate of 0.3 hours per minute (the product of the estimated 30 minute processing time to develop a fire mission and the same average fire development rate of 0.6 targets developed per minute). Transitional periods (e.g., start or end of operations) or other non-steady state operations can generate ratios that vary over time.

Table 1. KVA Productivity Ratios

		Weaponized UAV			
		Predator	Reaper	Sky Warrior	X-47B
Subprocess Productivity	Acquire targets	377	377	377	377
	Fire support coordination	189	189	189	189
	Fire mission development	943	3122	1222	3962
	Move weapons	50	23	44	607
	Engage targets	5094	70761	15212	254736
	Battlefield assessment	377	377	377	377
	Weapon	705	907	954	1067

⁷⁰ Although a single operational environment was simulated for this research, multiple and different environments can be simulated. Examples of characteristics of the operational scenario that can be elaborated to include dynamic variation in the entering target rate, distributions of target characteristics, and more target characteristics.

Note that, as described above, these productivities are ratios of accumulated learning time divided by accumulated processing time. Therefore, they are relative values. As expected, the three productivities for the subprocesses that are not impacted by the characteristics of the vehicle (Acquire targets Fire support coordination, and Battlefield assessment) do not change. These subprocesses are not impacted by different vehicles because the subprocess is the same for all of these vehicles. The application of system dynamics and KVA to the Analysis of Alternatives of other system alternatives such as improved logistics or vehicle technology used for recognizing and indentifying targets would generate changes in these KVA productivities. However, three important subprocesses that do impact total product productivity (“Weapons” row in Table 1) do vary (Fire mission development, Move weapons, and Engage targets).

Some of the ratios in Table 1 are relatively large when compared to returns on investment experienced in many industries, especially for the engage targets subprocess. A primary reason is that the numerator of these ratios includes the benefits of the technologies incorporated into the UAV for target engagement purposes. These technologies are extremely complex, are reflected in very large learning-time hours that are accumulated each time a target is engaged, and, therefore, generate high productivity ratios. Similarly, the denominator of these ratios reflects the time required to perform the subprocess, e.g., engage a target after it has been acquired, fire support coordinated, fire mission developed, and the UAV moved to station. Actual engagement times are relatively short for these UAV, further increasing the KVA productivity ratios for the engage target subprocess. Differences in learning-times across the UAV reflect their relative performance (e.g., automation of subprocesses previously performed by humans) and technologies are the primary causes of differences in the ratios across UAV in Table 1. Therefore, it is reasonable that the very large benefits of the under-development X-47B with its extremely advanced technology generate the largest ratios. Improved estimates of learning-times and processing times can improve the accuracy of these ratios. However, comparing the KVA productivity ratios for the engage targets subprocess with the ratios for the move weapons subprocess that is simpler (lower numerator) and takes longer (larger denominator) indicates that the rank order of the ratios reflects the relative returns of the different subprocesses.

Based on these and additional tests, the model is considered useful for the investigation of the integration of system dynamics and KVA.

Using System Dynamics and KVA to Improve AoA

Consider the following hypothetical example of the use of an integrated system dynamics/KVA model to improve the productivity estimates supporting an Analysis of Alternatives. Assume that a new version of the Predator UAV is being developed to enable it to engage opposing UAVs. Due to the much higher speeds and agility of UAVs compared to most land-based targets, the fraction of targets missed is expected to be higher than that currently experienced with the Predator. The acquisition program management team has access to some, but limited, resources (e.g., money, expert developer time until required delivery, technology development capabilities, approvals) to improve performance. Different stakeholders value payload, dash speed, and range differently and want the program management to recommend different improvements. Therefore, program management expects a rigorous review of its Analysis of Alternatives process and the results that will recommend one (and only one) of the improvements. As part of justification of the AoA decision, stakeholders of the two solutions not recommended are certain to require explanations of how and how much the recommended improvement impacts operational performance compared to the improvements that were not recommended. Cost would, most



likely, be their primary economic consideration as evidenced by the earlier case study examples. However, our analysis will focus on value compared to cost in terms of the capabilities of the systems.

Many alternatives have been proposed and are being considered. A few examples are⁷¹ as follows:

- **Increase the size of the power plant**, which can be used to increase the vehicle's payload, dash speed, or a combination of both. This requires an increase in fuel capacity to not reduce range.
- **Redesign the transmission**, which will increase the vehicle's dash speed.
- **Increase the fuel tank size**, which will increase the vehicle's range but decrease its dash speed unless the power plant is also increased.
- **Reduce the time required at base** between trips to station, which increases the time that the vehicle is on station and available for missions.

Performing detailed analysis of all the possible alternatives, such as by building and testing prototypes or very detailed simulations, often exceeds the resources of acquisition programs. Therefore, program managers are faced with the challenge of reducing a long list of potential alternatives to those that should definitely be included in the program, those that should be investigated further for potential inclusion, and those that should be rejected. The integration of system dynamics and KVA provided a timely and inexpensive means of evaluating all potential alternatives and reducing the "long list" of potential alternatives to a "short list" to be pursued or investigated further based on an objective and justifiable process. To do this, first the operation of the system with each potential alternative is simulated and the simulation model is used to calculate the KVA productivity ratios for the subprocesses and system as a whole. Table 2 provides an example of a portion of such an analysis for the hypothetical upgrading of the Predator UAV using the model described above.

⁷¹ There are interdependencies and tradeoffs in these alternatives, such as needing to increase the power plant size to maintain a given dash speed if the fuel tank size is increased. These are ignored here for simplicity. However, in application to an actual program developers would describe specific sets of features (e.g., possible versions of a vehicle) for analysis.

Table 2. Predator UAV Upgrade Program KVA Productivity Ratios for Analysis of Alternatives

		<u>Subprocess KVA ratios</u>			<u>Weapon System</u>	
		Develop fire mission	Move weapons	Engage targets	KVA ratio	% Change from Base Case
	Predator Base Case	943	50	5,094	705	0.0%
Improvement Alternative	Increase fuel capacity 100%	1,886	50	5,094	951	<u>34.9%</u>
	Increase fuel capacity 50%	1,415	50	5,094	831	17.9%
	Increase Power plant 100% for payload	849	50	7,641	771	9.4%
	Increase Power plant 50% for payload	849	50	7,641	771	9.4%
	Redesign transmission for 100% faster dash speed	943	100	10,188	741	5.1%
	Redesign transmission for 50% faster dash speed	943	75	7,641	727	3.1%
	Increase Power plant 100% for dash speed	849	100	10,188	717	1.7%
	Increase Power plant 50% for dash speed	849	75	7,641	702	-0.4%
	Reduce time at base 50%	943	52	5,094	699	-0.9%
	Reduce time at base 100%	943	51	5,094	695	-1.4%

The KVA productivity ratios are repetitive for some subprocesses across alternatives. This is partially because some alternatives do not change the impact on some subprocesses and partially because of the limited number of system interactions incorporated into this proof of concept model. However, the KVA+SD modeling results are adequate to show how more accurate results might be used in an AoA of these potential capabilities upgrades. Based on the results above, a program manager can assess the relative value-added of the eleven alternatives (including no-change as reflected by the Base Case) analyzed. Comparison to the base case (e.g., the existing vehicle in the case of the Predator upgrade) provides an estimate of relative performance improvement. Sorting the improvements provided by potential alternatives in decreasing order (Table 2) lists alternatives from most attractive (Increase fuel capacity 100%) to least attractive (Reduce time at base 50%). The AoA suggests that, if adequate resources are available, the alternative that improves the system the most is to increase the fuel capacity 100% because it improves the development of the fire missions. If inadequate resources are available to implement this alternative, then the program should attempt to increase the fuel capacity by 50% for similar reasons. The

program manager can also delete reducing time at the base and a 50% increase in power plant capacity that is used to increase dash speed from consideration since they do not improve performance. Certainly, other factors must be incorporated into a complete AoA (most notably development costs), but the results of the KVA analysis using the system dynamics model provide valuable information for making final recommendations.

Discussion and Conclusions

Summary

The Knowledge Value Added (KVA) approach to including benefits in Analysis of Alternatives (AoA) was integrated with a system dynamics model of weapon systems operations to investigate the potential of their integration to improve the accuracy of KVA productivity ratios and, thereby, AoA. An integrated model was developed for a generic mobile weapons system and calibrated to four existing weaponized Unmanned Aerial Vehicles (UAV). Six basic subprocesses of operations using the weapons were included in the simulation. KVA productivity ratios for each subprocess for each UAV were calculated, compared, and used to explain how the simulation and KVA approach work together to generate quantitative assessments of the relative value-added of each subprocess and whole weapon systems. A hypothetical upgrade program to one of the UAV was also simulated to demonstrate how the integrated model can be used to evaluate alternative upgrades and justify AoA decisions.

Evaluation of Results

An Analysis of Alternatives based on an integrated system dynamics/KVA model provides program management teams with several kinds of valuable information.

- **Quantified Measures of Improvement that include Benefits:** Measures of subprocesses and the weapon system as a whole are quantified using a common set of assumptions and values (those incorporated into the simulation model). Therefore, differences in ratios and the implied relative value of different alternatives are due to the differences in the alternatives themselves.
- **Overall System Improvement Estimates:** The weapon (versus subprocess) - productivity ratios reflect changes in total product operations. If adequate resources are available to adopt at least one alternative, then a list of alternatives ranked by overall system improvement (e.g., Table 2) can be used to “triage” alternatives into those that should definitely be pursued, those that require more investigation before deciding, and those that should be abandoned. For example, based on the right-hand column, Table 2 suggests that increasing the fuel capacity should be pursued before redesigning the transmission and that reducing the time at the base should not be considered further.
- **Guidance for Alternative Selection:** The analysis specifically identifies which alternatives improve which subprocesses and the whole weapon system and by how much. For example, Table 2 suggests that increasing fuel capacity increases the Fire mission development subprocess most and three alternatives significantly improve the engage target subprocess most.
- **Justification of Analysis of Alternatives Decisions:** When used with the simulation model, the KVA productivity ratios can help explain and justify Analysis of Alternatives decisions by providing a means of describing how each alternative impacts operations, subprocesses, and performance. For example, in the UAV case above, increasing power plant size increases the payload, which



increases the Payload/Lethal payload ratio, which increases the probability of destruction if hit ($p(\text{kill})$), which decreases the return flow “Incomplete kill rate” from the Battlefield Assessment Backlog to the Target engagement backlog (Figure 3). This reduces the average number of times that a target must be engaged to be destroyed, thereby improving the productivity of the engage target subprocess.

- **Guidance for Further Investigation:** In addition to suggesting better and worse alternatives to pursue, an integrated system dynamics/KVA model can provide guidance for further investigation of alternatives by indicating which subprocesses each alternative improves. For example, Table 2 indicates that the *reason* increasing fuel capacity improves performance is it improves the Fire mission development subprocess. The model (Figure 3) indicates that this occurs by increasing the vehicle range, which reduces the likelihood of losing a target if it is missed with ordinance. Acquisition program managers can use this information to focus further investigation and development of this alternative on Fire mission development to assure that these improvements in the specific operations identified with the model are realized during the alternative's development.

It is important to note that neither a system dynamics model nor a KVA analysis of this system alone can reasonably produce these results. Only by integrating system dynamics and the KVA approach are the benefits above available. Based on the modeling and assessment above, we conclude that integrated system dynamics/KVA models can significantly improve the Analysis of Alternatives and, thereby, acquisition.

Implications for Practice

The current work indicates that acquisition can be improved by using integrated system dynamics/KVA models in the Analysis of Alternatives. The rigorous development and use of integrated system dynamics/KVA models can have important implications for acquisition practice, including:

- The number of alternatives that can be analyzed with KVA can be increased due to the relative ease of reflecting alternatives in the operations simulation model compared to manually developing forecasts for use in KVA analysis. This increases the likelihood of identifying and selecting the optimal alternative.
- Justifications of AoA decisions can become stronger due to program managers having the ability to causally trace from specific potential and selected alternatives through their impacts on specific subprocesses and operations to performance.
- Justifications of AoA decisions can become more robust because they can reflect an analysis of a wider range of alternatives and more alternatives.
- Results of AoA can become more consistent through the use of a single, integrated model of system operations and KVA metrics instead of separate operations and value-added models.
- System dynamics/KVA models may be used to baseline product performance during the acquisition process. Performance of the product can be tracked over time and used to improve the model and thereby performance forecasts and AoA later in acquisition.

- ***Program management will select better alternatives*** due to the implications for practice above. This will generate more effective and potentially cheaper materiel solutions.

In addition, improving AoA and acquisition through integrated system dynamics/KVA models can improve CONOPS. The Javelin case study described previously in this report provides a vivid example of the ability of acquisition in general and improved AoA such as through integrated system dynamics/KVA modeling to impact tactics and strategy. Upon receipt and use of Javelin, operators expressed surprise that its range was twice that of the weapon it replaced. That increased range initiated improvements to tactics, techniques, and procedures (ttp) such as the use of Javelin to detonate Improvised Explosive Devices. This, in turn, can generate changes to strategies. Accurate forecasts of product subprocess performance (e.g., accuracy at longer range) can be used to plan CONOPS improvements before product delivery.

It is important to note that the purpose of the simulations of operations developed and illustrated in the current work is to capture the relative benefits and costs of different materiel alternatives, not to simulate the impacts of operations on opposing forces. The usefulness of models can only be judged in relation to the specific purpose for which they are built (Stermann, 2000). Therefore, because the purpose of integrated system dynamics and KVA models is to improve AoA, those models should be developed, assessed, and used separately from force-on-force and other simulations of operations developed for other purposes.

Future Work

The current proof-of-concept work has demonstrated the potential of integrated system dynamics/KVA models to improve Analysis of Alternatives and acquisition. Additional research can extend this work toward implementation and expanded application.

Opportunities include:

- Modeling a specific acquisition program in support of its Analysis of Alternatives process can develop and demonstrate the capability of operationalizing the approach tested here.
- If important uncertainties in system operations are incorporated into the system dynamics model it can be used to generate distributions of KVA productivities. These can be used to estimate the volatilities used in real options analysis, which has been demonstrated to be useful in DoD acquisition.
- The application of integrated system dynamics/KVA modeling to DoD product life cycle management can be investigated by using the model to generate forecasts of performance and KVA ratios during acquisition, comparing those forecasts with actual operations, and using the results to improve the model fidelity with the system. The improved model can then be used to analyze proposed changes or replacement of the system throughout its life cycle.

The Analysis of Alternatives is a particularly challenging part of DoD acquisition. Integrating system dynamics modeling and the Knowledge Value Added approach has been shown to be capable of improving that analysis and, thereby, alternative selection. Adapting this approach can significantly change and improve DoD acquisition practice.



References

- Cayas, C.P. (2007, August 3). Northrop wins Navy UCAS bid over Boeing. *Navy Times*. Retrieved March 18, 2010, from http://www.navytimes.com/news/2007/08/defense_ucas_northrop_070801/
- Flood, R.L., & Jackson, M.C. (1991). *Creative problem solving: Total systems intervention*. Chichester, UK: Wiley.
- Ford, D.N., & Dillard, J.T. (2008, May 14-15). Modeling the integration of open systems and evolutionary acquisition in DoD programs. In *Proceedings of the Fifth Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School.
- Ford, D.N., & Dillard, J.T. (2009, July). Modeling the performance and risks of evolutionary acquisition. *Defense Acquisition Review Journal*, 16(2), 143-158.
- Ford, D.N., & Dillard, J.T. (2009, May 13-14). Modeling open architecture and evolutionary acquisition in ARCI with applications to RCIP. In *Proceedings of the Sixth Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School.
- Forrester, J.W. (1961). *Industrial dynamics*. Portland, OR: Productivity Press.
- Global Security. (2010, March 22). *Unmanned Aerial Vehicles (UAVs)*. Retrieved from <http://www.globalsecurity.org/intell/systems/uav-intro.htm>
- Jackson, M.C. (2003). *Systems thinking: Creative holism for managers*. Chichester, UK: Wiley.
- Jones, C.A. (1997). *Unmanned Aerial Vehicles (UAVs), an assessment of historical operations and future possibilities*. Research Department, Ari Command and Staff College.
- Lane, D.C., & Jackson, M.C. (1995). Only connect! An annotated bibliography reflecting the breadth and diversity of systems thinking. *Systems Research*, 12, 217-228.
- Northrup Grumman. (2010). *X-47B Pegasus UCAV-N*. Retrieved March 15, 2010, from <http://www.globalsecurity.org/military/systems/aircraft/x-47b.htm>
- Northrup Grumman. (2010). *X-47B UCAS*. Retrieved March 18, 2010, from <http://www.as.northropgrumman.com/products/nucasx47b/index.html>
- Sterman, J.S. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. NY: Irwin McGraw-Hill.
- Williams, A.J. (2001, May). Air support functionality in AFATDS—advance field artillery tactical data system. *FA Journal*. Retrieved from http://findarticles.com/p/articles/mi_m0IAU/is_3_6/ai_76610396/



PEO-IWS ACB Insertion Portfolio Optimization⁷²

Johnathan Mun, Tom Housel and Mark D. Wessman

Overview

Program Executive Office–Integrated Warfare Systems (PEO-IWS) engaged a team from the Naval Postgraduate School (NPS) to conduct a pilot study to apply the Knowledge Valued Added + Real Options + Integrated Risk Management + Portfolio Optimization (KVA + RO + IRM + PO) method to estimate the value stream created by the capabilities to be inserted within the Aegis Weapons System (AWS) through the Advanced Capability Build (ACB) process—as described in the PEO-IWS Surface Combat System Acquisition Management Plan (AMP)—given budget constraints and ship industrial availability schedules. The goal was to determine what order of capability insertion provided the best returns within an optimized portfolio, treating each capability as a real option. The KVA + RO + IRM + PO approach was used to estimate the warfighter value delivered by each capability within the context of a portfolio-optimization, integrated risk-management model. The results provide a set of options based on selected constraints for insertion of the capabilities over the period of interest (Fiscal Years 2014–2025) based on an optimized portfolio model. For detailed information on the KVA + RO + IRM + PO approach, see the technical appendix to this report.

The pilot study analysis articulated a notional value measure of military capability for a specified set of 23 capabilities to be considered, and examined four discrete approaches: (1) a ranking by value within a constrained total integration budget (optimal on budget), (2) a similar ranking by value and budget with a risk constraint added (optimal cost-risk), (3) ranking by value constrained by integration budget, with the additional constraint that a particular capability must be included in the first increment (Capability 2 must-have), and (4) the portfolio with the specified capability in the first increment and a risk constraint added (Capability 2 cost-risk). Each approach generated a distinct recommendation for the composition and sequencing of the capabilities within the ACB schedule. Under the ACB model, capabilities will be inserted within the AWS system every two years. ACBs are identified by the fiscal year in which the first ship receives the software upgrade. The analysis encompassed ACB 14 (2014) through ACB 18 (2018). The analysis included all the ships that would receive the ACBs for each two-year period from 2014–2025, and accounted for the ships being phased into the program through scheduled repair availability periods. This period was selected because at the end of that time, all ships with the AWS would have been inducted into the process. The analysis assumes that value would begin to accrue for a given ship as soon as an ACB was implemented in that ship and would continue to accrue through subsequent ACBs throughout the service life of the ship. ACBs beyond ACB 18 were not considered for the pilot analysis, but both additional capabilities and future ACBs can easily be added to the analysis. Within the ACB process, a ship

⁷² The analyses performed (Monte Carlo risk simulation, dynamic optimization, and real options analysis) apply the Risk Simulator software and ROV Modeler software tools available from Real Options Valuation, Inc. (www.realoptionsvaluation.com), and the software screen shots were reprinted with their permission. Although there exist several commercial off-the-shelf software products available for running optimization, Risk Simulator and ROV Modeler were the only tools found to be suitable due to their ability in handling real options analysis, stochastic optimization, risk simulation, and other requirements in the analyses performed. These software tools were developed by the author, Dr. Johnathan Mun.



receives the current ACB, as well as the previous ACBs not yet completed. For example, a ship entering the program with ACB 18 also receives the capabilities from ACB 14 and ACB 16 at the same time. A ship that receives the first ACB as ACB 14 will receive its next update with ACB 18 but will receive the ACB 16 capabilities at the same time. Aggregation of the value for individual ships provides a measure in terms of capability-ship-years.

Within the context of the KVA approach, the study assumed that a relatively objective and extensible metric for military value was the relative complexity of the software modules that implement each capability in the open architected AWS, and that the complexity of the component could be represented by the relative magnitude of the number of delivered source lines of code (DSLOC), given that the programming languages used were comparable. A second measure of relative military value was achieved through collection of subject-matter expert (SME) rankings of relative component complexity and mission criticality. The study methodology aggregated the component data to the capability level using SME mapping of components to capabilities. These measures correlated highly with the DSLOC rankings, providing a validation of the assumption that relative magnitude of DSLOC can provide a measure of military value. SME estimates of the complexity of code and DSLOC were presumed to be those of the capability components themselves (versus simply the integration) since warfighting functionality (the military value of the system) was implemented by the component. Integrating the components/capabilities into the system made the warfighting functions of the components/capabilities available to the user.

The cost basis for each insertion employed for the study was based on an aggregated average of high-, medium-, and low-cost estimates to integrate the required components into the AWS for given capability insertions. Cost of integration was used as the key cost parameter to provide the analysis based on PEO-IWS 1's perspective as the integrating agent, versus as a component provider. The correlations among the subjective measure of military value and relative complexity of components derived from judgments of SMEs from PEO-IWS 1.0 and the Naval Surface Warfare Center Dahlgren Division (NSWCDD) were very high, indicating that the estimates were reliable (Table 1). The correlation between costs for the insertions and value described using this method was very low, indicating that integration cost does not predict military value. These findings support the need for using the KVA method to determine expected military value (EMV), the objective measure used in portfolio optimization and selection.

The toolset applied—using real options, portfolio optimization, and integrated risk management—provides a means for quickly estimating the effects of various capability insertions over the period of interest. It provides management with the flexibility to examine various ACB capability insertion options given budget and ship availability constraints. The analysis for this study employed the following steps:

1. Data collection and analysis to determine the best proxy for Expected Military Value (EMV), using objective data on DSLOC, subject-matter estimates of complexity and mission criticality for each capability, and OPNAV (Chief of Naval Operations staff) sponsor and technical community priorities for each capability.
2. Static and dynamic optimization runs based on four different EMV measures for multi-criteria optimization, to determine the best allocation and selection of capabilities given a nominal \$150 million budget constraint for each ACB, using a range of cost estimates for integration provided by SMEs.



3. Combination of the four EMV methods to obtain the portfolio of options, results and recommendations for sequencing capability insertions for ACB 14, ACB 16, and ACB 18.
4. Computation of aggregate EMV values through the insertions of ACBs using actual planned ship-availability schedules as published in the Surface Ship Acquisition Management Plan (AMP).
5. Monte Carlo risk simulation of cost estimates to determine risk of budget and cost overruns.
6. Generation of an alternate scenario by applying OPNAV's Priority 1 capability (Capability 2) as a "must-have" in the portfolio selection, and identification of what other capabilities should be inserted in such a scenario and how this would affect accrued value over the ACB insertion timeframe.
7. Determination of a Portfolio Efficient Frontier, in which we determine multiple scenarios of increasing budget (i.e., if the \$150 million budget were increased to \$200 million, or \$250 million, or \$300 million, etc., what will the optimal portfolio look like for each budget; what capabilities should be added or replaced, and what would be their impacts on EMV?).
8. Repetition of the previous analyses—with an additional constraint that the portfolio selected must have an 85% or greater probability of completing within budget.

As follow-on to the work documented in this report, PEO-IWS should consider the following:

1. Strategic Real Options or analysis of alternatives, examining various courses of action should certain capabilities be linked or nested with respect to another (e.g., there might be a "platform capability" that might have a high initial cost but bring significant downstream options for add-on capabilities with significant EMV. Or, there may be mutually exclusive or dependent capabilities—with which the implementation of capability precludes another from being implemented or requires another to be implemented, or will reduce the cost and increase the total EMV of another capability when they are implemented together).
2. Additional modeling, such as adding new capabilities to the list, adding considerations of additional risk factors (e.g., technical, schedule).
3. Training and software implementation for risk simulation and optimization.

Adoption of the foregoing will lead to a more refined and robust analysis of the value, risk, and cost of future options for capability insertions for the Aegis system. The remainder of the report is sequenced as follows:

- Statement of Work (SOW) Objectives
- Problem Formulation
- Methodology
 - Data Collection
 - Portfolio Optimization
 - Expected Military Value



- Sequencing of ACB 14, ACB 16, ACB 18
- Alternate Scenario: Capability 2 Must-Have
- Aggregate EMV
- Monte Carlo Risk Simulation
- Portfolio Optimization's Efficient Frontier
- Optimization with a Risk Constraint

Statement of Work Objectives

The focus of this work is to conduct a pilot study to provide return on investment and real options/portfolio optimization analysis to help articulate value proposition in selection of capabilities for inclusion in the modernization of the Aegis Weapons System (AWS) in US Navy cruisers and destroyers. The analysis was to be demonstrated in a manner that would support the next budget submission cycle. In addition, the project was to use the Knowledge Value Added + Real Options + Integrated Risk Management + Portfolio Optimization (KVA + RO + IRM + PO) methodology, with supporting software, to aid in the process performance analysis and option-value estimation. The customer selected the processes and systems for the analysis to establish the baseline return on investment (ROI) estimates. This project focused on conducting the KVA + RO + IRM + PO analysis on the identified ACB insertion options by working with PEO-IWS 1 and NSWCCD personnel to establish the necessary baselines and analyses and, concurrently, to lay the foundation for developing the level of knowledge necessary for the organization to use and maintain the toolset going forward. This approach ensures that the managers of the process have a decision toolset and the knowledge to interpret the results of the analysis outputs. These tools include the applications of risk analysis, forecasting, risk hedging and management strategies, strategic real-options applications, project portfolio optimization and selection, and other related analytics. In addition, aggregate numbers used to support the building of a business case to meet the acquisition community requirements for the selected problem space were also to be documented. Management-level reports were provided to evaluate ongoing OA acquisition initiatives. The products of the pilot were developed in a manner that can provide a basis for extension and implementation across the PEO as a method and toolset to be used on an ongoing basis. This extended use will provide the ability to better manage acquisition decisions and to make the case for those decisions to both sponsors and the Acquisition chain of command.

Problem Formulation

The US Navy is constantly faced with many difficult portfolio optimization decisions. These decisions include allocating financial resources, building or expanding facilities and capabilities, and determining acquisition strategies. Such decisions might involve thousands or millions of potential alternatives. Considering and evaluating each of them would be impractical or even impossible. A model can provide valuable assistance in incorporating relevant variables when analyzing decisions and finding the best solutions for making decisions. Models capture the most important features of a problem and present them in a form that is easy to interpret. Models often provide insights that intuition alone cannot. An optimization model has three major elements: *decision variables*, *constraints*, and *objectives*. The optimization methodology finds the best combination or permutation of decision variables (e.g., which strategies to pursue and which projects to execute) in every



conceivable way such that the objective is maximized (e.g., return on investment, military value-added, proxies for revenues and income) or minimized (e.g., risk and costs) while still satisfying the constraints (e.g., budget and resources).

In order to obtain optimal values, one generally must search in an iterative or ad hoc fashion. This search involves running one iteration for an initial set of values, analyzing the results, changing one or more values, rerunning the model, and repeating the process until finding a satisfactory solution. This process can be very tedious and time-consuming, even for small models, and often it is not clear how to adjust the values from one iteration to the next. A more rigorous method systematically enumerates all possible alternatives. This approach guarantees optimal solutions if the model is correctly specified. If an optimization model depends on only two decision variables, and if each variable has 10 possible values, then trying each combination requires 100 iterations (10^2 alternatives). If each iteration is very short (e.g., two seconds), then the entire process could be done in approximately three minutes of computer time. However, instead of two decision variables, if the option set includes “go” or “no-go” decisions on 23 alternative selections—as in the case of the current ACB analysis—then trying all combinations requires 2.58×10^{22} iterations of alternatives. It is easily possible for complete enumeration to take months or even years to carry out on a supercomputer. Practicality, then, demands that the analyst employ some advanced algorithms and techniques in Risk Simulator and ROV Modeler for running the portfolio selection and optimization. Before embarking on solving an optimization problem, it is vital to understand the terminology of optimization—the terms used to describe certain attributes of the optimization process. These words include *decision variables*, *constraints*, and *objectives*.

Decision Variables are quantities over which the decision-makers have control; for example, the amount of a product to make, the number of dollars to allocate among different investments, or which projects to select from among a limited set. As an example, portfolio optimization analysis includes a “go” or “no-go” decision on particular projects. In addition, the dollar or percentage budget allocation across multiple projects also can be structured as decision variables.

Constraints describe relationships among decision variables that restrict the values of the decision variables. For example, a constraint might ensure that the total amount of money allocated among various investments cannot exceed a specified amount or, at most, one project from a certain group can be selected. Constraints also include budget and timing restrictions, minimum returns, or risk-tolerance levels.

Objectives give a mathematical representation of the model's desired outcome—such as maximizing EMV, benefits, and profit, or minimizing cost and risk—in terms of the decision variables. In financial analysis, for example, the objective may be to maximize returns while minimizing risks (maximizing the Sharpe's ratio or returns-to-risk ratio).

The solution to an optimization model provides a set of values for the decision variables that optimizes (maximizes or minimizes) the associated objective. If the real business conditions were simple, and if the future were predictable, then all data in an optimization model would be constant, making the model deterministic. In many cases, however, a deterministic optimization model cannot capture all the relevant intricacies of a practical decision-making environment. When a model's data are uncertain and can only be described probabilistically, the objective will have some probability distribution for any chosen set of decision variables. The analyst can find this probability distribution by simulating the model using Risk Simulator. An optimization model under uncertainty has several additional elements, including assumptions and forecasts.



Assumptions capture the uncertainty of model data using probability distributions, whereas forecasts are the frequency distributions of possible results for the model. Forecast statistics are summary values of a forecast distribution, such as the mean, standard deviation, and variance. The optimization process controls the optimization by maximizing or minimizing the objective. Each optimization model has one objective, a variable that mathematically represents the model's objective in terms of the assumption and decision variables. Optimization's job is to find the optimal (minimum or maximum) value of the objective by selecting and improving different values for the decision variables. When model data are uncertain and can only be described using probability distributions, the objective itself will have some probability distribution for any set of decision variables. In the current project's optimization analysis, the problem formulation is to optimize the Aegis ACB composition based on:

- Potential return on capability (return on investment, expected military value, and other multiple criteria),
- Investment constraints (e.g., \$150 million per ACB cycle),
- Ship schedule and availability, and
- Selecting the best combinations and permutations of capabilities using the portfolio optimization approach as a series of options.

Methodology

In this section, we discuss the methodology employed in more detail, with particular emphasis on the high-level understanding of the approach and the results. For the technical mathematical constructs, please refer to the Appendix for additional technical background readings. Briefly, the methodology employed is divided into several steps, as covered in the following subsections.

Data Collection

Data collection and analysis is the first step employed to determine the best proxy for Expected Military Value (EMV) and cost estimates of each capability. To that end, we relied on data ranging from objective values such as delivered source lines of code (DSLOC) of software, semi-objective measures such as estimates of integration cost for each capability (using high, most-likely, and low estimates for cost, so that we can perform a Monte Carlo risk simulation later), to more subjective estimates from subject-matter experts (SMEs) on the amount of functional complexity and operational criticality for each component. PEO-IWS representatives also provided OPNAV and acquisition community priorities for each capability. The analysis demonstrated that complexity is proportionate to value, but there were low correlations between EMV and cost estimates—indicating that we cannot reliably use cost alone as an estimate to determine the best portfolio allocation for maximizing EMV. The correlation matrix is shown in Table1.

Table 1. Correlation Matrix

Linear Correlation Matrix	Cost Midpoint	SME Mean Value Added	DSLOC Complexity	Technical Priority H-L	OPNAV Priority H-L	EMV: SMEVA and DSLOC	EMV: All Weighted & Common Sized	EMV: OPNAV & Technical	EMV: OPNAV & DSLOC
Cost Midpoint	1								
SME Mean Value Added	0.07	1							
DSLOC Complexity	0.40	0.84	1						
Technical Priority H-L	0.46	0.42	0.42	1					
OPNAV Priority H-L	0.34	0.49	0.43	0.89	1				
EMV: SMEVA and DSLOC	0.39	0.86	1.00	0.42	0.44	1			
EMV: All, Weight, Sized	0.36	0.91	0.93	0.65	0.69	0.94	1		
EMV: OPNAV & Technical	0.41	0.47	0.44	0.97	0.97	0.44	0.69	1	
EMV: OPNAV & DSLOC	0.44	0.82	0.90	0.72	0.78	0.90	0.98	0.77	1

Risk and uncertainty can also be estimated based on various criteria (in the current analysis, we use cost uncertainty as a proxy for risk), and value is assumed to be generated as capabilities are realized through installation in specific ships over time. To get started with the data collection, we had to perform the following steps:

- Establish operational definitions of value and cost of each ACB capability insertion.
- Identify the projected ship schedule to establish availability for ACB insertions every two years.
- Obtain SME identification and description of ACB components and capabilities.
- Undertake model generation and iterations with various inputs, including running cross-correlations to determine the impact and validity of SME estimates.

The data sources used include:

- AMP Ver. 5.4 (27 Oct 2008) documentation of moving to an OA approach in ACB insertions,
- Ship schedule, capability candidates for integration, components of the system to be changed (mapped to the capabilities), and integration cost for each capability provided by IWS 1,
- SMEs estimates of complexity and mission criticality for components, and
- DSLOC for each component.

Figure 1 shows a sample of the collected data used in the analysis. The analysis began by assuming 23 capabilities (more can be added later as required). Next, the analysis applied the average of the SME value-added estimates; the high, most-likely, and low cost estimates; the OPNAV and technical priorities; and the DSLOC for each component, as specified in the *Surface Combat System Objective Architecture*. Using these raw variables, we generated various EMV metrics by accounting for the SME mean value-added estimates and DSLOC, and weighting them—as well as common-sizing their mean values—to determine a comprehensive metric, considering OPNAV and technical priority only, and combining OPNAV priority with DSLOC estimates. Clearly, other metrics can be easily applied in the model if required.

Common Size Factor				Common sizing inputs and using weights to obtain EMV				Multicriteria Objectives				OPNAV Weight		50%					
Weighting Scheme				35%				2.00				2.00		SMEVA and DSLOC Only		All Weighted & Common Sized		OPNAV Priority & DSLOC	
Capability	High	Mid	Low	SME Mean Value Added	Technical Priority	OPNAV Priority	DSLOC Complexity	Cost Simulation	Technical Priority H-L	OPNAV Priority H-L	EMV Score	EMV Score	EMV Score	EMV Score	Objective Used	CV			
Capability 1	Data Redacted	Data Redacted	Data Redacted	43.00	3	2	278	Data Redacted	21	22	321.00	37.68	43.00	35.90	37.68	0.1834			
Capability 2				28.00	2	1	541		22	23	569.00	42.24	45.00	50.05	42.24	0.0903			
Capability 3				13.00	1	3	58		23	21	71.00	19.78	44.00	23.90	19.78	0.0719			
Capability 4				40.33	5	4	635		19	20	675.33	48.04	39.00	51.75	48.04	0.0576			
Capability 5				17.67	7	7	134		17	17	151.67	21.07	34.00	23.70	21.07	0.0407			
Capability 6				35.67	4	6	392		20	18	427.67	37.60	38.00	37.60	37.60	0.1568			
Capability 7				50.67	6	5	549		18	19	599.67	48.05	37.00	46.45	48.05	0.0481			
Capability 8				47.67	10	11	675		14	13	722.67	48.41	27.00	46.75	48.41	0.0689			
Capability 9				19.00	8	8	109		16	16	128.00	20.07	32.00	21.45	20.07	0.0382			
Capability 10				33.67	11	9	189		13	15	222.67	26.80	28.00	24.45	26.80	0.0603			
Capability 11				20.00	9	10	88		15	14	108.00	18.78	29.00	18.40	18.78	0.0707			
Capability 12				27.67	14	12	159		10	12	106.67	21.85	22.00	19.95	21.85	0.1137			
Capability 13				27.67	12	13	159		12	11	186.67	22.15	23.00	18.95	22.15	0.0408			
Capability 14				44.67	13	14	523		11	10	567.67	40.24	21.00	36.15	40.24	0.0956			
Capability 15				17.67	20	15	134		4	9	151.67	14.77	13.00	15.70	14.77	0.0603			
Capability 16				17.67	21	16	134		3	8	151.67	14.17	11.00	14.70	14.17	0.0816			
Capability 17				34.33	19	17	328		5	7	362.33	27.10	12.00	23.40	27.10	0.0454			
Capability 18				17.67	22	18	134		2	6	151.67	13.27	8.00	12.70	13.27	0.0372			
Capability 19				14.00	23	19	81		1	5	95.00	9.54	6.00	9.05	9.54	0.0340			
Capability 20				17.67	15	20	134		9	4	151.67	14.77	13.00	10.70	14.77	0.0164			
Capability 21				22.67	18	21	144		6	3	166.67	15.67	9.00	10.20	15.67	0.0454			
Capability 22				11.67	16	22	180		8	2	191.67	13.38	10.00	11.00	13.38	0.0468			
Capability 23				10.00	17	23	95		7	1	105.00	9.23	8.00	5.75	9.23	0.0603			

Starting with 23 capabilities, more to be added later

High, likely, low values for running risk simulations

Technical and OPNAV priorities

Complexity modeled using DSLOC

Intermediate computations: risk simulation assumptions, re-adjusted priorities, expected military score, cost-based risk coefficients

Figure 1. Input Assumptions

Portfolio Optimization

Static and dynamic optimization runs were executed based on four different EMV measures for multi-criteria optimization to determine the best allocation and selection of capabilities given a nominal \$150 million budget constraint for each ACB. Figure 2 shows the portfolio optimization model, in which we have the 23 capabilities listed (clearly, we can add as many additional capabilities as required, as long as we have valid data and assumptions for each capability). The EMV values (column C in Figure 2) show the value of the composite EMV metric, depending on the calculation method chosen, the EMV value using risk-simulated cost estimates (column D), and EMV value using an estimate of risk (columns E and F) for each component (for the initial study, we impute the risk as the budget cost overrun and variability, whereas we can add additional variables and measures of risk later, as required). Finally, there is a column of decision variables, or “go” and “no-go” variables (column H), which are the decisions that are being optimized, such that the total portfolio EMV objective (cell C28) is maximized. The total cost of the portfolio is also computed (cell D28), and the portfolio is run subject to a cost constraint of less than or equal to \$150 million (cell D29). For future applications, we can add to the existing optimization model by also considering:

- Additional Capabilities as required, beyond the initial list of 23,
- Optimization and selection of Components, instead of Capabilities,
- New and alternate EMV metrics beyond the four EMV estimates currently used,
- Additions of cross-constraints such as mutually exclusive projects and capabilities, and the dependence of one capability on another,

- Inclusion of additional constraints such as full-time equivalences, facilities, etc., and
- Estimates of technical or schedule risk.

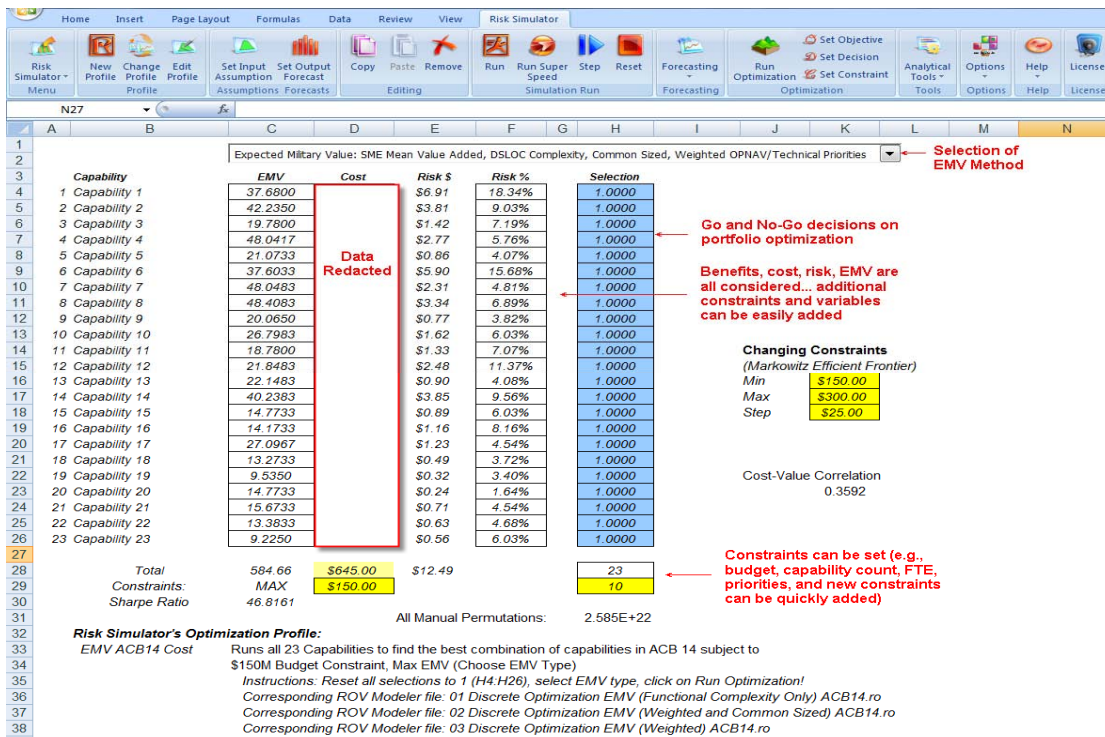


Figure 2. Portfolio Optimization Model

Figure 3 illustrates the portfolio optimization and capability selection setup using the Risk Simulator software. It shows Static and Dynamic Optimization routines run on multiple decision variables and constraints. It also shows the exact specifications of the model.

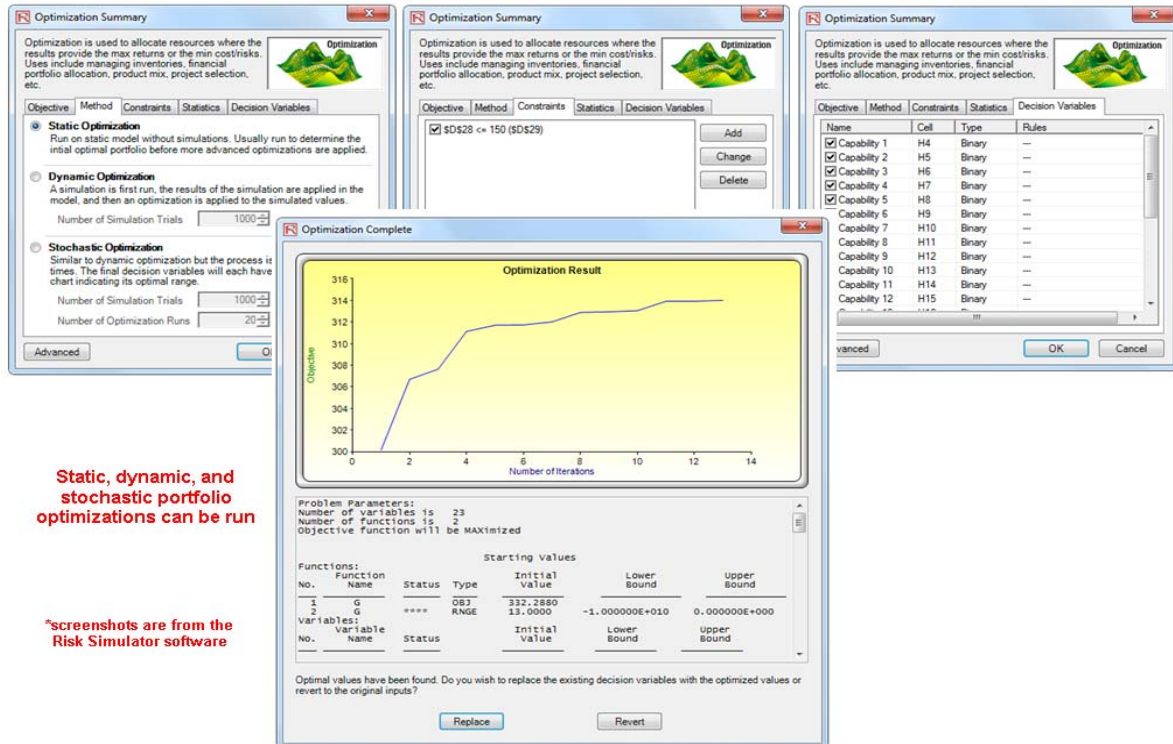


Figure 3. Optimization Model Setup in Risk Simulator
Expected Military Value (EMV)

The next step in the analysis was to apply the combination of 4 EMV methods to obtain the portfolio of options results and recommendations for sequencing capability insertions for ACB 14, ACB 16, and ACB 18. Figures 4 through 7 show the details of each portfolio optimization run and their corresponding capability set selected. The specification of each optimization depends on the EMV that is selected. For instance, the first set of results below is run based on using the mean of the subject-matter experts' (SME) value-added estimates, software lines of code (DSLOC), and OPNAV and technical priorities, and all these variables are combined through weighting and common-sizing the averages. The portfolio optimization is run to determine the best capabilities to select to maximize the total EMV for the portfolio, while at the same time maximizing the EMV, subject to the \$150 million budget constraint. These results indicate a multi-criteria optimization routine, in which various objectives or EMVs are used in the portfolio-selection process.

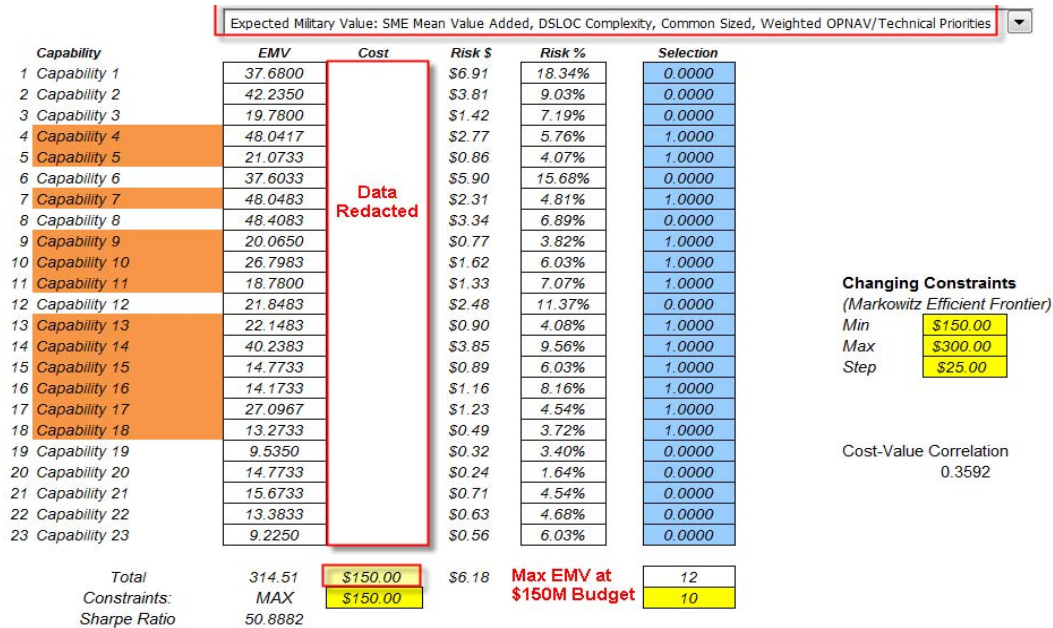


Figure 4. ACB 14 Optimization Results

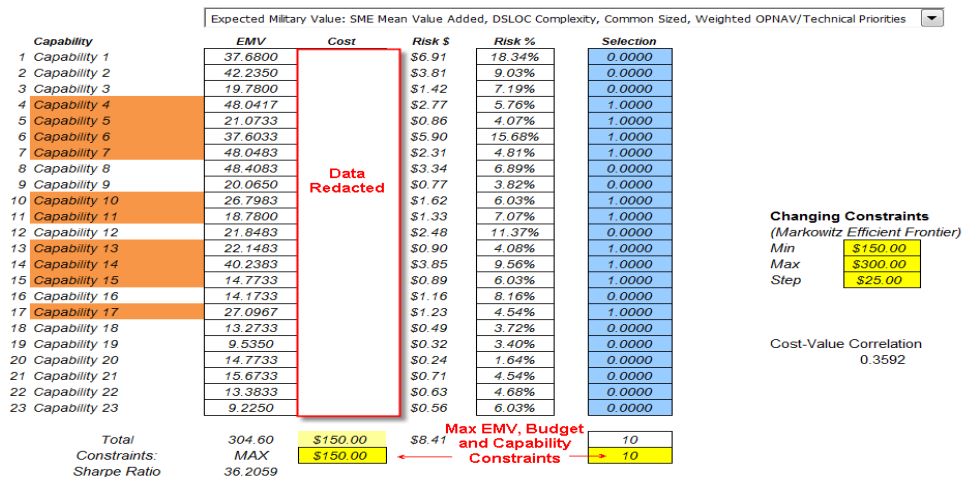


Figure 5. ACB 14 Optimization Results II



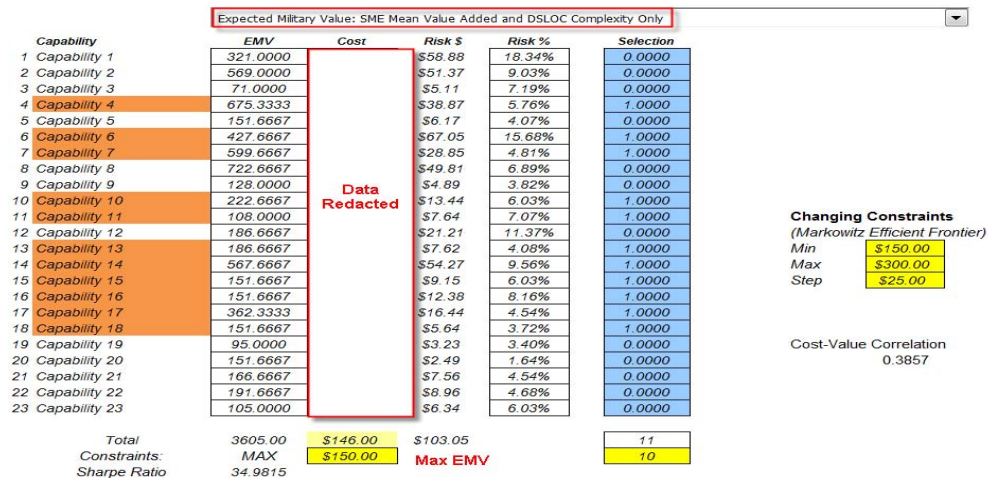


Figure 6. ACB 14 Optimization Results III

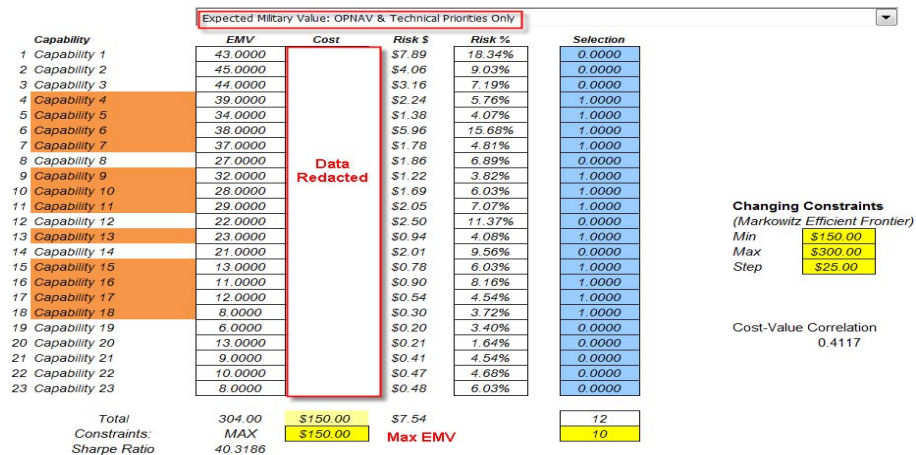


Figure 7. ACB 14 Optimization Results IV

Figure 8 summarizes the capabilities chosen based on each of the four EMV approaches, and the resultant recommendations for implementation. Specifically, the last column shows the optimal decision based on a portfolio of options of decisions. For instance, the following capabilities should be considered as optimal for ACB 14: Capabilities 4, 7, 10, 11, 13, 14, 15, 16, 17, and 18. Using the multi-criteria optimization, the analysis does not simply rely on a single estimate of EMV, but is able to employ the data from multiple facets and triangulate the best course of action. Figure 8 shows the results. It illustrates that certain components are always selected regardless of the EMV metric used, providing a higher level of comfort in the analysis, and these are the components we recommend (last column in the figure). Further, there are multiple components that 3 out of 4 of the optimization routines suggest executing; in most cases, these are not considered the top 10 components, but nonetheless important in the current ACB 14. These 4 EMV choices provide a view on the Analysis of Alternatives. The results of this analysis postulates which components are considered the top 10 and which are not, while still being critical in the ACB 14 portfolio. Using these four EMV options, we have four optimal portfolios, and we can quickly determine the best Course of Action (shown as the last column in Figure 8).

Discrete Optimization EMV (OPNAV and Technical Priorities Only) ACB14							Recommendation Option 1 \$134	Recommendation Option 2 146
Constraint Capability	SMEMVA, DSLOC, Priorities, Sized \$150M	SMEMVA, DSLOC, Priorities, Sized \$150M Top 10	SMEMVA, DSLOC \$150M	Technical & OPNAV Priorities \$150M	OPNAV Priority & DSLOC \$150M			
Capability 1								
Capability 2								
Capability 3								
Capability 4	GO	GO	GO	GO	GO		1	1
Capability 5	GO	GO		GO	GO		1	
Capability 6		GO	GO	GO	GO			1
Capability 7	GO	GO	GO	GO	GO		1	1
Capability 8								
Capability 9	GO			GO	GO			
Capability 10	GO	GO	GO	GO	GO		1	1
Capability 11	GO	GO	GO	GO	GO		1	1
Capability 12								
Capability 13	GO	GO	GO	GO	GO		1	1
Capability 14	GO	GO	GO				1	1
Capability 15	GO	GO	GO	GO	GO		1	1
Capability 16	GO		GO	GO	GO		1	1
Capability 17	GO	GO	GO	GO	GO		1	1
Capability 18	GO		GO	GO	GO		1	1
Capability 19								
Capability 20								
Capability 21								
Capability 22								
Capability 23								

Figure 8. ACB 14 Portfolio of Options Results

Sequencing of ACB 14, ACB 16, ACB 18

The next step in the analysis was performing a sequential compound portfolio option by examining ACB 16 and ACB 18. In other words, based on the analysis for ACB1 4, the budget of no more than \$150 million will be spent on 11 capabilities, with the remaining 12 capabilities still available for future execution. So, with another \$150 million budget in ACB 16, the portfolio optimization was rerun with the truncated list of available capabilities, and the results shown in Figure 9 were obtained. The analysis considers capabilities that may not be available until later ACBs by simply including them in the process beginning with the earliest ACB for which they are ready for integration. The optimization is repeated based on each of the four multi-criteria objectives and provides a list of recommended capabilities to execute (highlighted box in Figure 10 shows the recommended components in this ACB).

Expected Military Value: SME Mean Value Added, DSLOC Complexity, Common Sized, Weighted OPNAV/Technical Priorities						
Capability	EMV	Cost	Risk \$	Risk %	Selection	
1 Capability 1	37.6800	Data Redacted	\$6.91	18.34%	1.0000	ACB16
2 Capability 2	42.2350		\$3.81	9.03%	0.0000	
3 Capability 3	19.7800		\$1.42	7.19%	0.0000	
4 Capability 4	48.0417		\$2.77	5.76%		ACB14
5 Capability 5	21.0733		\$0.86	4.07%	1.0000	ACB16
6 Capability 6	37.6033		\$5.90	15.68%		ACB14
7 Capability 7	48.0483		\$2.31	4.81%		ACB14
8 Capability 8	48.4083		\$3.34	6.89%	0.0000	
9 Capability 9	20.0650		\$0.77	3.82%	1.0000	ACB16
10 Capability 10	26.7983		\$1.62	6.03%		ACB14
11 Capability 11	18.7800		\$1.33	7.07%		ACB14
12 Capability 12	21.8483		\$2.48	11.37%	1.0000	ACB16
13 Capability 13	22.1483		\$0.90	4.08%		ACB14
14 Capability 14	40.2383		\$3.85	9.56%		ACB14
15 Capability 15	14.7733		\$0.89	6.03%		ACB14
16 Capability 16	14.1733		\$1.16	8.16%		ACB14
17 Capability 17	27.0967		\$1.23	4.54%		ACB14
18 Capability 18	13.2733		\$0.49	3.72%		ACB14
19 Capability 19	9.5350		\$0.32	3.40%	1.0000	ACB16
20 Capability 20	14.7733		\$0.24	1.64%	1.0000	ACB16
21 Capability 21	15.6733		\$0.71	4.54%	1.0000	ACB16
22 Capability 22	13.3833		\$0.63	4.68%		ACB14
23 Capability 23	9.2250		\$0.56	6.03%	1.0000	ACB16
Total	149.87	\$141.00	\$7.50		8	
Constraints:		\$150.00			10	
Sharpe Ratio	19.9844					

Figure 9. ACB 16 Optimization Results

Discrete Optimization EMV ACB16				Recommendation Option 1 \$131
SMEMVA, DSLOC, Priorities, Sized \$150M	SMEMVA, DSLOC \$150M	Technical & OPNAV Priorities \$150M	OPNAV Priority & DSLOC \$150M	
GO	GO	GO	GO	1
ACB14	ACB14	ACB15	ACB16	
GO		GO	GO	1
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
	GO			
GO		GO	GO	1
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
GO		GO	GO	1
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
GO		GO	GO	1
GO	GO	GO	GO	1
GO	GO	GO	GO	1
GO	GO	GO	GO	1

Figure 10. ACB 16 Portfolio of Options Results

The analysis continued with the portfolio optimization on ACB 18. The highlighted boxes in Figures 11 and 12 show the recommended component to execute in ACB 18, and in this case it is Capability 2. Capability 2 holds the highest OPNAV priority, but when the analysis includes the other inputs into the EMV metrics (DSLOC, Technical Priority, Weights, and SME Estimated Value-Add), and considers the high cost (\$126 million) with respect to the allowed portfolio budget (\$150 million), Capability 2 fails the selection for ACB 14 and ACB 16, and is only recommended in ACB 18. However, the analysis can consider alternate scenarios in which OPNAV Priority is taken as the most important criteria, and Capability 2 is specified as a “must-have” in ACB 14. The analysis was then rerun to determine the optimal portfolio given this new requirement. The results of that run are documented in the next section. The report next illustrates the effects on EMV of selecting Capability 2 in ACB 14 through to the year 2025. The report then continues with the Efficient Frontier analysis to show what additional components should be added in each ACB if additional budget is allocated (e.g., what if the budget were extended to \$175 million or \$200 million, and so forth, to determine at what point perhaps more critical components would have been selected). The optimization analysis is highly flexible to accommodate such alternate scenarios and requirements.

Expected Military Value: SME Mean Value Added, DSLOC Complexity, Common Sized, Weighted OPNAV/Technical Priorities						
Capability	EMV	Cost	Risk \$	Risk %	Selection	
1 Capability 1	37.6800	Data Redacted	\$6.91	18.34%		ACB16
2 Capability 2	42.2350		\$3.81	9.03%	1.0000	ACB18
3 Capability 3	19.7800		\$1.42	7.19%	0.0000	
4 Capability 4	48.0417		\$2.77	5.76%		ACB14
5 Capability 5	21.0733		\$0.86	4.07%		ACB16
6 Capability 6	37.6033		\$5.90	15.68%		ACB14
7 Capability 7	48.0483		\$2.31	4.81%		ACB14
8 Capability 8	48.4083		\$3.34	6.89%	0.0000	
9 Capability 9	20.0650		\$0.77	3.82%		ACB16
10 Capability 10	26.7983		\$1.62	6.03%		ACB14
11 Capability 11	18.7800		\$1.33	7.07%		ACB14
12 Capability 12	21.8483		\$2.48	11.37%		ACB16
13 Capability 13	22.1483		\$0.90	4.08%		ACB14
14 Capability 14	40.2383		\$3.85	9.56%		ACB14
15 Capability 15	14.7733		\$0.89	6.03%		ACB14
16 Capability 16	14.1733		\$1.16	8.16%		ACB14
17 Capability 17	27.0967		\$1.23	4.54%		ACB14
18 Capability 18	13.2733		\$0.49	3.72%		ACB14
19 Capability 19	9.5350		\$0.32	3.40%		ACB16
20 Capability 20	14.7733		\$0.24	1.64%		ACB16
21 Capability 21	15.6733		\$0.71	4.54%		ACB16
22 Capability 22	13.3833		\$0.63	4.68%		ACB14
23 Capability 23	9.2250		\$0.56	6.03%		ACB16
Total	42.24	\$126.00	\$3.81		1	
Constraints:	MAX	\$150.00			10	
Sharpe Ratio	11.0759					



Figure 11. ACB 18 Optimization Results

Discrete Optimization EMV ACB18				Recommendation Option 1 \$126
SMEMVA, DSLOC, Priorities, Sized \$150M	SMEMVA, DSLOC \$150M	Technical & OPNAV Priorities \$150M	OPNAV Priority & DSLOC \$150M	
ACB16	ACB16	ACB16 GO	ACB16 GO	1
ACB14	ACB14	ACB14	ACB14	
ACB16	ACB16	ACB16	ACB16	
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
GO	GO			
ACB16	ACB16	ACB16	ACB16	
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
ACB16	ACB16	ACB16	ACB16	
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
ACB14	ACB14	ACB14	ACB14	
ACB16	ACB16	ACB16	ACB16	
ACB16	ACB16	ACB16	ACB16	
ACB16	ACB16	ACB16	ACB16	
ACB16	ACB16	ACB16	ACB16	

Figure 12. ACB 18 Portfolio of Options

Alternate Scenario: Capability 2 “Must-Have”

The optimization model and approach used is highly adaptive and flexible. The next analytic step conducted was the specification of one or more components as a "must have," specifically, Capability 2 was set as a mandatory capability for inclusion in ACB 14. Figure 13, illustrates the generation of an alternate scenario by applying OPNAV's Priority 1 on Capability 2 as a “must-have” in the portfolio selection and by identifying what other capabilities should be inserted in each ACB examined in such a scenario. When the integration budget is constrained to \$150 million, if a significant portion of it is allocated to Capability 2, then only a little is left over for other components. Figure 13 shows what these components are (i.e., on the last column of Figure 13, the decision variable set as 1 indicates a “go” decision, whereas 0 indicates a “no-go”).

Expected Military Value: SME Mean Value Added, DSLOC Complexity, Common Sized, Weighted OPNAV/Technical Priorities					
Capability	EMV	Cost	Risk \$	Risk %	Selection
1 Capability 1	37.6800	\$6.91	\$6.91	18.34%	0.0000
2 Capability 2	42.2350	\$3.81	\$3.81	9.03%	1.0000
3 Capability 3	19.7800	\$1.42	\$1.42	7.19%	0.0000
4 Capability 4	48.0417	\$2.77	\$2.77	5.76%	0.0000
5 Capability 5	21.0733	\$0.86	\$0.86	4.07%	0.0000
6 Capability 6	37.6033	\$5.90	\$5.90	15.68%	0.0000
7 Capability 7	48.0483	\$2.31	\$2.31	4.81%	0.0000
8 Capability 8	48.4083	\$3.34	\$3.34	6.89%	0.0000
9 Capability 9	20.0650	\$0.77	\$0.77	3.82%	0.0000
10 Capability 10	26.7983	\$1.62	\$1.62	6.03%	0.0000
11 Capability 11	18.7800	\$1.33	\$1.33	7.07%	1.0000
12 Capability 12	21.8483	\$2.48	\$2.48	11.37%	0.0000
13 Capability 13	22.1483	\$0.90	\$0.90	4.06%	0.0000
14 Capability 14	40.2383	\$3.85	\$3.85	9.56%	0.0000
15 Capability 15	14.7733	\$0.89	\$0.89	6.03%	0.0000
16 Capability 16	14.1733	\$1.16	\$1.16	8.16%	1.0000
17 Capability 17	27.0967	\$1.23	\$1.23	4.54%	1.0000
18 Capability 18	13.2733	\$0.49	\$0.49	3.72%	1.0000
19 Capability 19	9.5350	\$0.32	\$0.32	3.40%	0.0000
20 Capability 20	14.7733	\$0.24	\$0.24	1.64%	0.0000
21 Capability 21	15.6733	\$0.71	\$0.71	4.54%	0.0000
22 Capability 22	13.3833	\$0.63	\$0.63	4.68%	0.0000
23 Capability 23	9.2250	\$0.56	\$0.56	6.03%	0.0000
Total	115.56	\$149.00	\$4.40		5
Constraints:		MAX			10
Sharpe Ratio	26.2374				

Figure 13. Capability 2 “Must-Have” in ACB 14



Aggregate EMV

By computing EMV accrued for each ACB, the analysis can then track the aggregation of value in terms of EMV available by ship-year based on the year of installation of each ACB in an individual ship. This approach permits representing the total capability available to the fleet as a single number. Figure 14 shows the ship availabilities for ACB insertions. Using this ship schedule and availability for ACB insertion and applying the optimal EMV values, the aggregate EMV values through the year 2025 become known (Figure 15). The second curve in Figure 15 demonstrates that EMV over time is marginally reduced by requiring Capability 2 to be included in ACB 14. Figure 15 also shows the “catch up” effect of the ACB process. Even though the introduction of a capability might be delayed from one ACB to the next, the total number of ships possessing the capability will become the same after the fourth ACB period if the delayed capability is included in the next update. An alternative to consider is to maintain the \$150 million budget across all ACBs, but at the same time increase the ACB 14 budget to include a “special insertion budget” to cover Capability 2 and maintain the portfolio as suggested previously.

	A	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	YEAR IN PORT FOR AEGIS UPGRADE BASED ON 4 YEAR REFRESH CYCLE (TOTALS FOR EACH ACB AT BOTTOM)													
2	ACB 14		15 ACB 16		17 ACB 18		19 ACB 20		21 ACB 22		23 ACB 24		25	
4	SHIP ID													
5	CG52 - 1	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4605.48
6	2	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4604.48
7	3	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4476.55
8	4	115.56	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4555.4
9	5	115.56	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4555.4
10	6	0	0	115.56	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	4043.88
11	7	0	115.56	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4043.88
12	8	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4604.48
13	9	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4604.48
14	10	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4604.48
15	11	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4604.48
16	12	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4604.48
17	13	0	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	4092.96
18	14	0	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	4092.96
19	15	0	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	4092.96
20	16	0	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	4092.96
21	17	0	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	4092.96
22	18	0	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	4092.96
23	19	115.56	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4554.4
24	20	115.56	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4554.4
25	21	115.56	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4554.4
26	22	115.56	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4554.4
27	23	115.56	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4554.4
28	24	115.56	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4554.4
29	25	0	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4042.88
30	26	0	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4042.88
31	27	0	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4042.88
32	28	0	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4042.88
33	29	0	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4042.88
34	30	0	115.56	115.56	115.56	511.52	511.52	511.52	511.52	511.52	511.52	511.52	511.52	4042.88
35	31	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4603.48
36	32	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4603.48
37	33	0	0	383.59	383.59	383.59	383.59	511.52	511.52	511.52	511.52	511.52	511.52	4603.48

Figure 14. ACB 14 Ship Availabilities

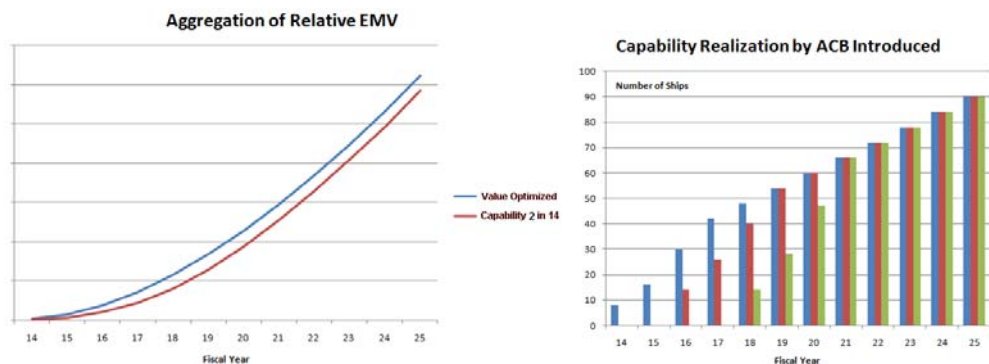


Figure 15. ACB 14 Aggregate EMV

Monte Carlo Risk Simulation

Monte Carlo Risk Simulation of cost estimates is used to determine the risk of budget and cost overruns (project timing overruns can also be determined if required). A sample of the simulated risk analysis results is shown in Figure 16. While capability selection is the key

question addressed in this study, the risk analysis results are necessary and support some of the optimization analysis. For future modeling and decision-analysis work, the proper determination of appropriate risk measures, potentially including cost, schedule, technology, and other risks, would be appropriate and beneficial. The analyst can model all these uncertainties using the Risk Simulator software tools. Instead of relying on single-point estimates for cost and scheduled completion times, distributions of cost and time through expert estimates, comparable historical data, and expectations of high, most-likely, and low estimates for each input should be employed. The analytic method then specifies that these values be simulated thousands of times with the software to generate all possible outcomes and scenarios, and the results are then interpreted to examine the risks inherent in each ACB insertion. Applying this method to the first analysis, the results indicate that although the expected total cost is \$150 million, there is 83.30% chance that the budget will be exceeded. In fact, to be 99% sure that there is sufficient money to cover the potential cost-creep, the budget would have to be increased to \$171 million, indicating the need of a \$21 million cushion. Similarly, the analyst can apply the methodology to determine the probability of the occurrence of schedule overruns. Other risks, such as technology risk, may be expressed in other ways to provide inputs to the simulator software.

Risk Simulation of Cost

*Screen shots from Risk Simulator software

Risk analysis and 100,000 simulation trials on cost estimations...

We can determine the probability that ACB-X will exceed the \$150M budget, determine what \$X will yield a 99% certainty of sufficient budget to cover all costs

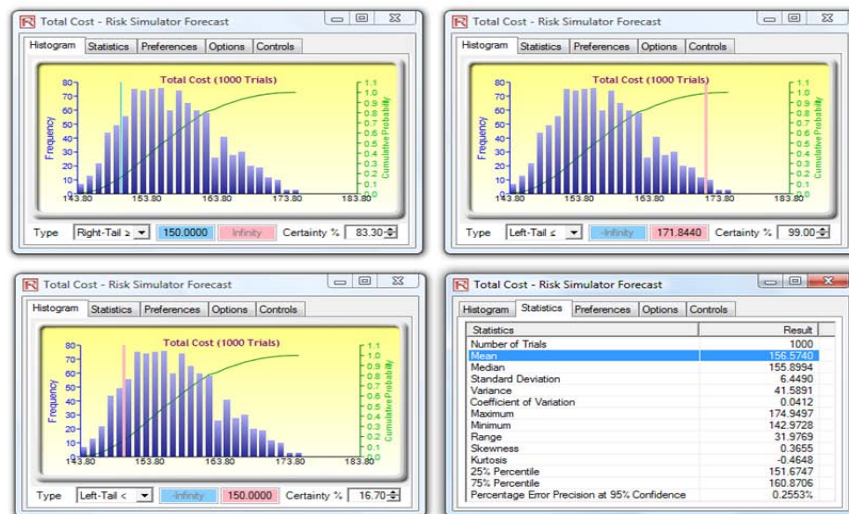


Figure 16. Risk Analysis on Cost for ACB 14

Portfolio Optimization's Efficient Frontier

The portfolio Efficient Frontier analysis determines multiple scenarios of increasing budget (i.e., if the \$150 million budget were increased to \$200 million, or \$250 million, or \$300 million, and so forth, in various increments, what will the optimal portfolio look like; what capabilities should be added or replaced; and what are the impacts on EMV?). This analysis provides useful input for deliberations with the sponsor early in the budget-development process and yields data-driven sets of alternatives for various levels of funding.

Running the optimization procedure yields an optimal portfolio of projects in which the constraints are satisfied. This represents a single optimal portfolio point on the Efficient Frontier—for example, Portfolio B on the Efficient Frontier chart in Figure 17. Then, by

subsequently changing some of the constraints—for instance, by increasing the budget—the analyst can rerun the optimization to produce another optimal portfolio given these new constraints. Therefore, a series of optimal portfolio allocations can be determined and graphed. This graphical representation of all optimal portfolios is called the portfolio's Efficient Frontier. At this juncture, each point represents a portfolio allocation. For instance, Portfolio B might represent capabilities 1, 2, 5, 6, 7, 8, 10, 15, and so forth, while Portfolio C might represent capabilities 2, 6, 7, 9, 12, 15, and so forth—each resulting in different EMV, tactical, military, or comprehensive scores, and portfolio returns. It is up to the decision-maker to decide which portfolio represents the best decision and if sufficient resources exist to execute these projects. Typically, in an Efficient Frontier analysis, a decision-maker would select projects for which the marginal increase in benefits is positive, and the slope is steep. In the next example, that decision-maker would rather select Portfolio D rather than Portfolio E, as the marginal increase is negative on the y-axis (e.g., EMV). That is, spending too much money may actually reduce the overall EMV; hence, this portfolio should not be selected. Also, in comparing Portfolios A and B, a decision-maker would be more inclined to choose B, as the slope is steep and the same increase in budget requirements (x-axis) would return a much higher percentage EMV (y-axis). The decision to choose between Portfolios C and D would depend on available resources, and the decision-maker must decide if the added benefits warrant and justify the added budget and costs. Figures 18 through 22 illustrate the results from the Efficient Frontier analysis by changing the budget constraint from \$150 million to \$300 million by incrementing it \$25 million in each step.

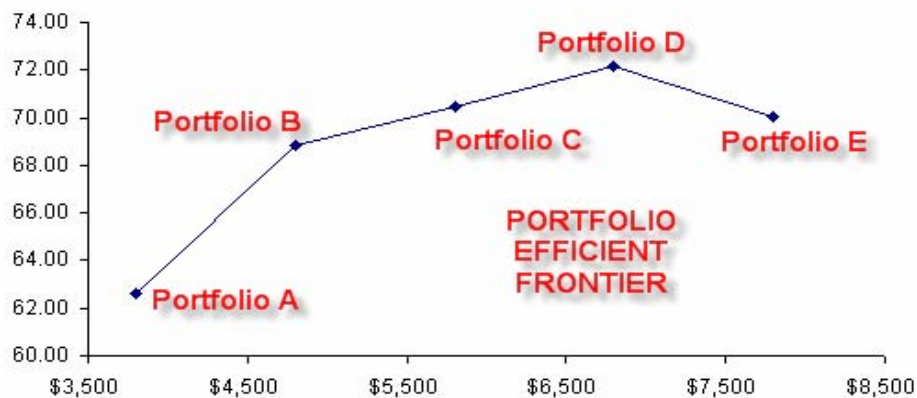
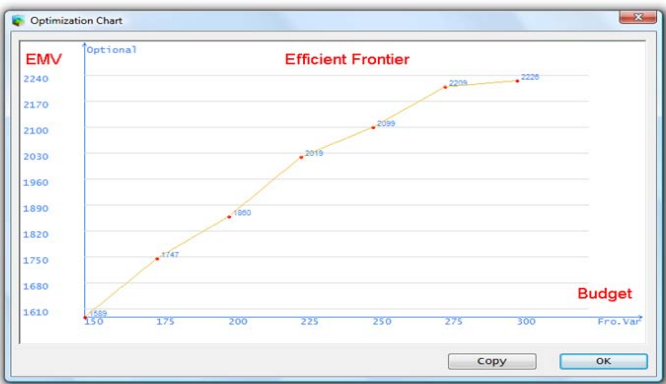


Figure 17. The Theory of Portfolio Efficient Frontier

Portfolio Optimization: Efficient Frontier Analysis



ACB 14 Markowitz efficient frontier (x-axis is budget amount, and y-axis is total expected military value)

*Screen shots from ROV Optimizer software

Figure 18. Portfolio Efficient Frontier

Discrete Optimization EMV (OPNAV and Technical Priorities Only) ACB14								
Constraint	\$150M	\$175M	\$200M	\$225M	\$250M	\$275M	\$300M	
Expected Cost	\$146M	\$173M	\$200M	\$224M	\$248M	\$274M	\$300M	
Capability 1	GO	GO	GO	GO	GO	GO	GO	
Capability 2	GO	GO	GO	GO	GO	GO	GO	
Capability 3	GO	GO	GO	GO	GO	GO	GO	
Capability 4	GO	GO	GO	GO	GO	GO	GO	
Capability 5	GO	GO	GO	GO	GO	GO	GO	
Capability 6	GO	GO	GO	GO	GO	GO	GO	
Capability 7	GO	GO	GO	GO	GO	GO	GO	
Capability 8	GO	GO	GO	GO	GO	GO	GO	
Capability 9	GO	GO	GO	GO	GO	GO	GO	
Capability 10	GO	GO	GO	GO	GO	GO	GO	
Capability 11	GO	GO	GO	GO	GO	GO	GO	
Capability 12	GO	GO	GO	GO	GO	GO	GO	
Capability 13	GO	GO	GO	GO	GO	GO	GO	
Capability 14	GO	GO	GO	GO	GO	GO	GO	
Capability 15	GO	GO	GO	GO	GO	GO	GO	
Capability 16	GO	GO	GO	GO	GO	GO	GO	
Capability 17	GO	GO	GO	GO	GO	GO	GO	
Capability 18	GO	GO	GO	GO	GO	GO	GO	
Capability 19	GO	GO	GO	GO	GO	GO	GO	
Capability 20	GO	GO	GO	GO	GO	GO	GO	
Capability 21	GO	GO	GO	GO	GO	GO	GO	
Capability 22	GO	GO	GO	GO	GO	GO	GO	
Capability 23	GO	GO	GO	GO	GO	GO	GO	

Figure 19. Efficient Frontier with Additional Budget Allocation I

Optimization EMV (SMEVVA DSLOC Weighted, Common Sized, Priorities) ACB14 Efficient								
Constraint	\$150M	\$175M	\$200M	\$225M	\$250M	\$275M	\$300M	
Capability	\$150M	\$175M	\$200M	\$225M	\$248M	\$274M	\$297M	
Capability 1	GO	GO	GO	GO	GO	GO	GO	
Capability 2	GO	GO	GO	GO	GO	GO	GO	
Capability 3	GO	GO	GO	GO	GO	GO	GO	
Capability 4	GO	GO	GO	GO	GO	GO	GO	
Capability 5	GO	GO	GO	GO	GO	GO	GO	
Capability 6	GO	GO	GO	GO	GO	GO	GO	
Capability 7	GO	GO	GO	GO	GO	GO	GO	
Capability 8	GO	GO	GO	GO	GO	GO	GO	
Capability 9	GO	GO	GO	GO	GO	GO	GO	
Capability 10	GO	GO	GO	GO	GO	GO	GO	
Capability 11	GO	GO	GO	GO	GO	GO	GO	
Capability 12	GO	GO	GO	GO	GO	GO	GO	
Capability 13	GO	GO	GO	GO	GO	GO	GO	
Capability 14	GO	GO	GO	GO	GO	GO	GO	
Capability 15	GO	GO	GO	GO	GO	GO	GO	
Capability 16	GO	GO	GO	GO	GO	GO	GO	
Capability 17	GO	GO	GO	GO	GO	GO	GO	
Capability 18	GO	GO	GO	GO	GO	GO	GO	
Capability 19	GO	GO	GO	GO	GO	GO	GO	
Capability 20	GO	GO	GO	GO	GO	GO	GO	
Capability 21	GO	GO	GO	GO	GO	GO	GO	
Capability 22	GO	GO	GO	GO	GO	GO	GO	
Capability 23	GO	GO	GO	GO	GO	GO	GO	

Figure 20. Efficient Frontier with Additional Budget Allocation II

Discrete Optimization EMV (SMEVVA DSLOC Only) ACB14 Efficient Frontier								
Constraint	\$150M	\$175M	\$200M	\$225M	\$250M	\$275M	\$300M	
Capability	\$146M	\$173M	\$200M	\$224M	\$248M	\$274M	\$297M	
Capability 1	GO	GO	GO	GO	GO	GO	GO	
Capability 2	GO	GO	GO	GO	GO	GO	GO	
Capability 3	GO	GO	GO	GO	GO	GO	GO	
Capability 4	GO	GO	GO	GO	GO	GO	GO	
Capability 5	GO	GO	GO	GO	GO	GO	GO	
Capability 6	GO	GO	GO	GO	GO	GO	GO	
Capability 7	GO	GO	GO	GO	GO	GO	GO	
Capability 8	GO	GO	GO	GO	GO	GO	GO	
Capability 9	GO	GO	GO	GO	GO	GO	GO	
Capability 10	GO	GO	GO	GO	GO	GO	GO	
Capability 11	GO	GO	GO	GO	GO	GO	GO	
Capability 12	GO	GO	GO	GO	GO	GO	GO	
Capability 13	GO	GO	GO	GO	GO	GO	GO	
Capability 14	GO	GO	GO	GO	GO	GO	GO	
Capability 15	GO	GO	GO	GO	GO	GO	GO	
Capability 16	GO	GO	GO	GO	GO	GO	GO	
Capability 17	GO	GO	GO	GO	GO	GO	GO	
Capability 18	GO	GO	GO	GO	GO	GO	GO	
Capability 19	GO	GO	GO	GO	GO	GO	GO	
Capability 20	GO	GO	GO	GO	GO	GO	GO	
Capability 21	GO	GO	GO	GO	GO	GO	GO	
Capability 22	GO	GO	GO	GO	GO	GO	GO	
Capability 23	GO	GO	GO	GO	GO	GO	GO	

Figure 21. Efficient Frontier with Additional Budget Allocation III



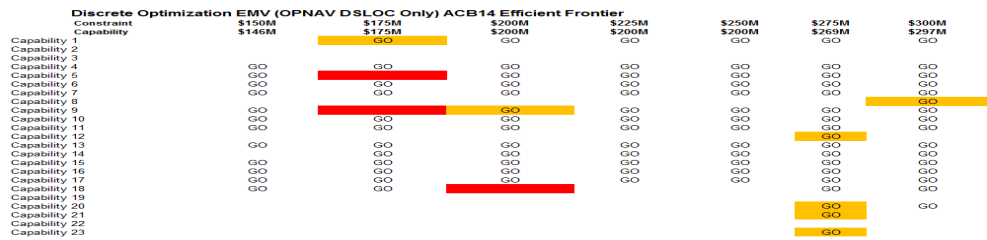


Figure 22. Efficient Frontier with Additional Budget Allocation IV
Optimization with a Risk Constraint

Figure 23 shows an example optimization run in which we can set cost as the stochastic constraint. That is, seeing that cost overruns typically occur in development, we can set the risk simulation optimization combination model such that we want a portfolio where there is a 90% probability that the \$150 million budget is not exceeded. In this sample run, we see that this can be accomplished by replacing Capability 9 with Capability 23, at a lower cost, thereby still creating the maximum EMV possible while maintaining a 90% probability that total portfolio cost will be under the required \$150 million budget constraint. Alternatively, as shown previously in Figure 16, if the optimal portfolio is still desirable, then the \$150 million budget needs to account for a potential overrun of \$22 million. That is, there is a 99% probability that the total portfolio budget will be under \$172 million. There is clearly a risk-value tradeoff occurring in this situation; the higher the probability of lower budget overruns, the lower the anticipated EMV. This tradeoff is also seen very clearly in the Efficient Frontier analysis, in which the results demonstrate that the higher the budget allocation, the greater the EMV. The final step in the analysis was a sample run with ACB 14, 16, and 18 based on the EMV applying all input assumptions, and with the additional contingent constraint the total budget used will not exceed \$150 million for each ACB for at least 85% of the time. The results are shown in Figures 24, 25, and 26 for each of the ACBs sequenced. The resulting budget overrun risks are depicted in Figure 27. When this constraint is applied, the analysis yields a different portfolio selection for each ACB, and the probability of **not** exceeding the requisite \$150 million budget becomes 97.99%, 90.90% and 99.99% for the three ACB years. Running the model with Capability 2 as mandatory in ACB 14 and applying the risk constraint, generates yet another set of selections, yielding confidence levels of not exceeding the budget of \$150 million. Figure 28 shows the resulting aggregate EMV across all ships, revised to include the risk-cost portfolio. It is clear that the opportunity cost of applying the risk constraint, while measurable, is minimal **for this case** and actually represents less reduction in EMV overall than does the mandatory selection of a capability.

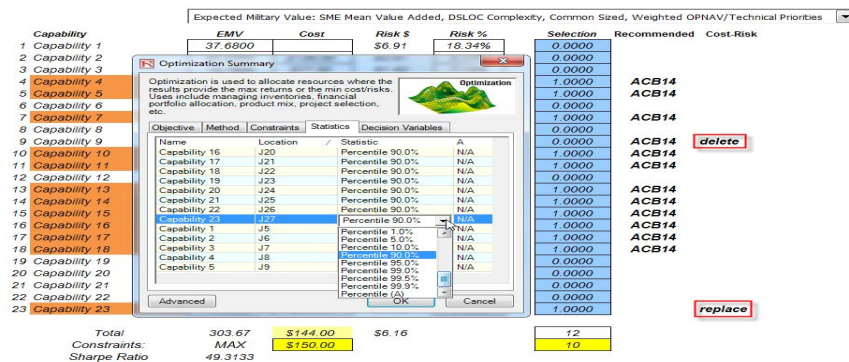


Figure 23. Cost-based Risk Optimization Alternative

Expected Military Value: SME Mean Value Added, DSLOC Complexity, Common Sized, Weighted OPNAV/Technical Priorities							
Capability	EMV	Cost	Risk \$	Risk %	Selection	Recommended	Cost-Risk
1 Capability 1	37.6800		\$6.91	18.34%	0.0000		
2 Capability 2	42.2350		\$3.81	9.03%	0.0000		
3 Capability 3	19.7800		\$1.42	7.19%	0.0000		
4 Capability 4	48.0417		\$2.77	5.76%	1.0000	ACB14	
5 Capability 5	21.0733		\$0.86	4.07%	1.0000	ACB14	
6 Capability 6	37.6033		\$5.90	15.68%	0.0000		
7 Capability 7	48.0483		\$2.31	4.81%	1.0000	ACB14	
8 Capability 8	48.4083		\$3.34	6.89%	0.0000		
9 Capability 9	20.0650		\$0.77	3.82%	1.0000	ACB14	
10 Capability 10	26.7983		\$1.62	6.03%	1.0000	ACB14	
11 Capability 11	18.7800		\$1.33	7.07%	1.0000	ACB14	
12 Capability 12	21.8483		\$2.48	11.37%	0.0000	ACB14	
13 Capability 13	22.1483		\$0.90	4.08%	1.0000	ACB14	
14 Capability 14	40.2383		\$3.85	9.56%	1.0000	ACB14	
15 Capability 15	14.7733		\$0.89	6.03%	0.0000		
16 Capability 16	14.1733		\$1.16	8.16%	1.0000	ACB14	
17 Capability 17	27.0967		\$1.23	4.54%	1.0000	ACB14	
18 Capability 18	13.2733		\$0.49	3.72%	1.0000	ACB14	
19 Capability 19	9.5350		\$0.32	3.40%	0.0000		
20 Capability 20	14.7733		\$0.24	1.64%	0.0000		
21 Capability 21	15.6733		\$0.71	4.54%	0.0000		
22 Capability 22	13.3833		\$0.63	4.68%	0.0000		
23 Capability 23	9.2250		\$0.66	6.03%	0.0000		
Total		299.74	\$140.00	\$6.12	11		
Constraints:			MAX	\$150.00	10		
Sharpe Ratio		49.0105					

Figure 24. Cost-based Risk Optimization for ACB 14

Expected Military Value: SME Mean Value Added, DSLOC Complexity, Common Sized, Weighted OPNAV/Technical Priorities									
Capability	EMV	Cost	Risk \$	Risk %	Selection	Recommended	Cost-Risk		
1 Capability 1	37.6800		\$6.91	18.34%	1.0000	ACB16	Cost Risk		
2 Capability 2	42.2350		\$3.81	9.03%	0.0000				
3 Capability 3	19.7800		\$1.42	7.19%	0.0000				
4 Capability 4	48.0417		\$2.77	5.76%					
5 Capability 5	21.0733		\$0.86	4.07%					
6 Capability 6	37.6033		\$5.90	15.68%	1.0000	ACB16	Cost Risk		
7 Capability 7	48.0483		\$2.31	4.81%					
8 Capability 8	48.4083		\$3.34	6.89%	0.0000				
9 Capability 9	20.0650		\$0.77	3.82%					
10 Capability 10	26.7983		\$1.62	6.03%					
11 Capability 11	18.7800		\$1.33	7.07%					
12 Capability 12	21.8483		\$2.48	11.37%	1.0000	ACB16	Cost Risk		
13 Capability 13	22.1483		\$0.90	4.08%					
14 Capability 14	40.2383		\$3.85	9.56%					
15 Capability 15	14.7733		\$0.89	6.03%	1.0000	ACB16	Cost Risk		
16 Capability 16	14.1733		\$1.16	8.16%					
17 Capability 17	27.0967		\$1.23	4.54%					
18 Capability 18	13.2733		\$0.49	3.72%					
19 Capability 19	9.5350		\$0.32	3.40%	0.0000				
20 Capability 20	14.7733		\$0.24	1.64%	1.0000	ACB16	Cost Risk		
21 Capability 21	15.6733		\$0.71	4.54%	1.0000	ACB16	Cost Risk		
22 Capability 22	13.3833		\$0.63	4.68%	0.0000				
23 Capability 23	9.2250		\$0.66	6.03%	1.0000	ACB16	Cost Risk		
Total		151.58	\$129.00	\$9.51		7			
Constraints:		MAX	\$150.00			10			
Sharpe Ratio		15.9455							

Figure 25. Cost-based Risk Optimization for ACB 16

Expected Military Value: SME Mean Value Added, DSLOC Complexity, Common Sized, Weighted OPNAV/Technical Priorities							
Capability	EMV	Cost	Risk \$	Risk %	Selection	Recommended	Cost Risk
1 Capability 1	37.6800	Data Redacted	\$6.91	18.34%	ACB16		
2 Capability 2	42.2350		\$3.81	9.03%	0.0000		
3 Capability 3	19.7800		\$1.42	7.19%	0.0000		
4 Capability 4	48.0417		\$2.77	5.76%	ACB14		
5 Capability 5	21.0733		\$0.86	4.07%	ACB14		
6 Capability 6	37.6033		\$5.90	15.68%	ACB16		
7 Capability 7	48.0483		\$2.31	4.81%	ACB14		
8 Capability 8	48.4083		\$3.34	6.89%	1.0000	ACB18	Cost Risk
9 Capability 9	20.0650		\$0.77	3.82%	ACB14		
10 Capability 10	26.7983		\$1.62	6.03%	ACB14		
11 Capability 11	18.7800	\$1.33	7.07%	ACB14			
12 Capability 12	21.8483	\$2.48	11.37%	ACB16			
13 Capability 13	22.1483	\$0.90	4.08%	ACB14			
14 Capability 14	40.2383	\$3.85	9.56%	ACB14			
15 Capability 15	14.7733	\$0.89	6.03%	ACB16			
16 Capability 16	14.1733	\$1.16	8.16%	ACB14			
17 Capability 17	27.0967	\$1.23	4.54%	ACB14			
18 Capability 18	13.2733	\$0.49	3.72%	ACB14			
19 Capability 19	9.5350	\$0.32	3.40%	ACB14	ACB18	Cost Risk	
20 Capability 20	14.7733	\$0.24	1.64%	ACB16			
21 Capability 21	15.6733	\$0.71	4.54%	ACB16			
22 Capability 22	13.3833	\$0.63	4.68%	0.0000			
23 Capability 23	9.2250	\$0.66	6.03%	ACB16			
Total:		57.94	\$95.00		2		
Constraints:		MAX	\$150.00		10		

Figure 26. Cost-based Risk Optimization for ACB 18

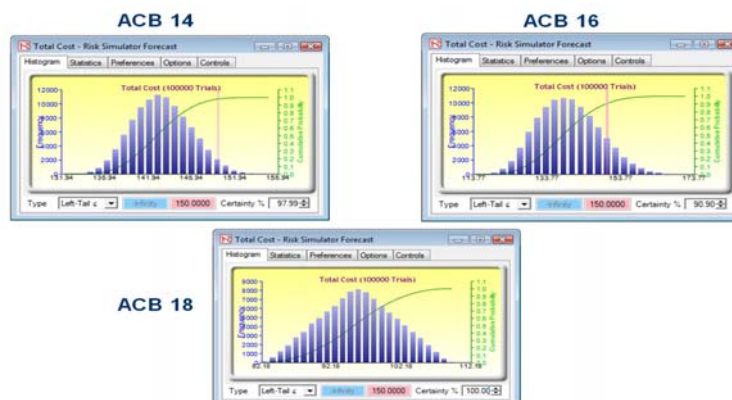


Figure 27. Cost-Risk Probabilities for ACB 14, 16, and 18

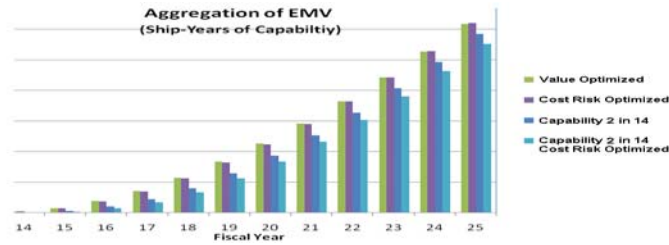


Figure 28. Aggregate EMV for 2014–2025

Figures 29 and 30 summarize the results of all portfolio runs as a simple visual matrix and list of summary statistics. The portfolio runs' visual matrix of results consists of four columns of portfolios: (1) optimization based on the required total budget of \$150 million, assuming all costs are exact and have no risk (Optimal on Budget); (2) cost with uncertainty and risk that there will be budget overruns—with an added constraint of a portfolio with no more than 15% probability of a budget overrun, or 85% probability or higher, that the total budget of \$150 million will not be exceeded (Optimal Cost-Risk); (3) Capability 2 is a required component in ACB 14 (Capability 2 Must-Have); and (4) Capability 2 as a Must-Have in ACB 14, with the added assumption of cost-risk as defined above (Capability 2 Cost-Risk). Similar ACB years are color-coded (e.g., green for ACB 14, blue for ACB 16, orange for ACB 18, and red for components that are not selected in these three ACBs, allowing for potential implementation later).

Figure 30 shows the summary key statistics of these portfolios, listing the number of capabilities implemented in each ACB cycle, the total expected budget used, the total EMV, and the probability that the ACB will be under budget. Clearly, the non-Cost-Risk portfolios bear higher total implementation costs with higher EMVs (high risk means high returns); however, the probability of being under budget is low, and the probability the budget will be exceeded is high. In this First Phase analysis, due to the analysis being run on only 23 capabilities, as expected, the EMVs are reallocated over time in various amounts (e.g., Capability 2 must-have will yield a smaller initial EMV due to the higher cost and moderate EMV value of executing Capability 2, but the catch-up happens in the subsequent ACBs). Figure 30 also examines the risk distribution of the budget based on the different portfolio criteria. For instance, we see that if we apply the value optimization without regard to the risk of cost overruns the median (or 50th percentile) budget is \$153.2 million, above the budget constraint of \$150 million. Alternatively, if we consider the risk of cost overruns, the median is only \$142.9 million, providing a buffer for any overruns. In fact, we see that the 85th percentile is \$146.6 million, and the 95th percentile is 148.7 million—both under the required \$150 million. Further, there is a 97.90% probability that this portfolio will come in under the \$150 million budget (for the sake of clarity, these values are highlighted in yellow in Figure 31). To reduce and hedge the risk of cost overruns—the expected budget used is less (\$139 million as opposed to \$146 million), with a return on EMV that is also less (310.98 as compared to 299.74). Therefore, to hedge and reduce the risk of cost overrun, the Navy spends less and gets less. This can be viewed as keeping some of the budget aside for the worst case scenario—therefore leaving less money available to invest in additional capabilities (the remaining statistics are fairly self-explanatory). One alternative to utilizing the highest number of capabilities, maximizing the EMV, and yet coming under budget, is to consider strategic real options in contract negotiations.

Decision-makers should exercise caution in the use of risk constraints to restrict consideration of portfolios. Blind selection of a risk-limited portfolio may result in excessive

opportunity cost if other means exist to reduce risk in the input data. For example, better cost estimates in one or more of the components would reduce volatility in that component and, thereby, make it less likely to be excluded due to a high contribution to aggregate risk. Similarly, altering the cost profile through risk mitigation efforts in the contract structure (caps, fixed-price provisions, etc.) changes the input and will change the output from the model. By applying the appropriate risk-mitigation measures and by rerunning the analysis, decision-makers may provide a better portfolio selection than simply constraining the analysis through applying a risk cap. Intelligent use of the toolset as a decision aid maximizes its value to the manager.

ACB 14 + ACB 16 + ACB 18

Capability	Optimal on Budget	Optimal Cost-Risk	Capability 2 Must-Have	Capability 2 Cost-Risk
Capability 1	ACB16	ACB16	ACB16	ACB18
Capability 2	ACB18	Later	ACB14	ACB14
Capability 3	Later	Later	Later	Later
Capability 4	ACB14	ACB14	ACB16	ACB16
Capability 5	ACB16	ACB14	ACB16	ACB16
Capability 6	ACB14	ACB16	ACB18	ACB18
Capability 7	ACB14	ACB14	ACB16	ACB16
Capability 8	Later	ACB18	ACB18	Later
Capability 9	ACB16	ACB14	ACB18	ACB16
Capability 10	ACB14	ACB14	ACB16	ACB16
Capability 11	ACB14	ACB14	ACB14	ACB14
Capability 12	ACB16	ACB16	ACB18	ACB18
Capability 13	ACB14	ACB14	ACB16	ACB16
Capability 14	ACB14	ACB14	ACB16	ACB16
Capability 15	ACB14	ACB16	ACB16	ACB18
Capability 16	ACB14	ACB14	ACB14	ACB16
Capability 17	ACB14	ACB14	ACB14	ACB16
Capability 18	ACB14	ACB14	ACB14	ACB16
Capability 19	ACB16	ACB18	Later	Later
Capability 20	ACB16	ACB16	Later	ACB18
Capability 21	ACB16	ACB16	Later	ACB18
Capability 22	Later	Later	Later	Later
Capability 23	ACB16	ACB16	ACB16	ACB18

Figure 29. Summary of All Sequenced and Optimized Portfolios

ACB 14 + ACB 16 + ACB 18
*rounded to the nearest 0.1

	Optimal on Budget	Optimal Cost-Risk	Capability 2 Must-Have	Capability 2 Cost-Risk
Total Capabilities ACB14	11	11	5	2
Total Capabilities ACB16	8	7	9	10
Total Capabilities ACB18	1	2	4	7
EMV ACB14	310.98	299.74	115.56	61.02
EMV ACB16	149.87	151.58	268.03	280.96
EMV ACB18	42.24	57.94	127.93	151.58
Total Cost ACB14	\$146.00	\$139.00	\$149.00	\$129.00
Total Cost ACB16	\$141.00	\$129.00	\$150.00	\$137.00
Total Cost ACB18	\$126.00	\$95.00	\$141.00	\$129.00
Total Spent on ACB14-18	\$413.00	\$363.00	\$440.00	\$395.00
Probability of Under Budget ACB14	29.70%	97.90%	41.50%	90.80%
Probability of Under Budget ACB16	72.23%	90.90%	16.25%	99.90%
Probability of Under Budget ACB18	94.80%	99.90%	72.90%	90.90%
ACB14 Median 50th Percentile on Budget	\$153.20	\$142.90	\$152.60	\$132.30
ACB14 Median 85th Percentile on Budget	\$160.22	\$146.60	\$166.60	\$146.30
ACB14 Median 95th Percentile on Budget	\$164.30	\$148.70	\$173.90	\$153.50
ACB16 Median 50th Percentile on Budget	\$145.40	\$137.90	\$156.50	\$139.50
ACB16 Median 85th Percentile on Budget	\$153.50	\$147.30	\$164.80	\$143.30
ACB16 Median 95th Percentile on Budget	\$157.80	\$152.70	\$169.20	\$145.30
ACB18 Median 50th Percentile on Budget	\$128.90	\$95.30	\$145.10	\$137.90
ACB18 Median 85th Percentile on Budget	\$142.90	\$101.30	\$153.50	\$147.30
ACB18 Median 95th Percentile on Budget	\$150.20	\$104.30	\$158.40	\$152.70

Figure 30. Summary of All Sequenced and Optimized Portfolios' Summary Statistics



THIS PAGE INTENTIONALLY LEFT BLANK



Panel #24 – Considerations for Developing and Nurturing the Future Acquisition Workforce

Thursday, May 13, 2010

3:30 p.m. –
5:00 p.m.

Chair: James Gill, Technical Advisor to the Director of Contracting, Space & Missile Systems Center

Determining the Optimal Staffing Level for the Acquisition Workforce: Different Models in Different Services Yield Different Results

Tim Reed, Naval Postgraduate School

Compensation, Culture and Contracts: The Realities of the DoD's Blended Workforce

Kathy Loudin, Defense Acquisition University

Industry Perceptions of Department of Defense Program Manager Competencies

Roy Wood, Defense Acquisition University



Determining the Optimal Staffing Level for the Acquisition Workforce: Different Models in Different Services Yield Different Results

Tim Reed—Timothy Reed is a visiting Professor at the Naval Postgraduate School, where he teaches graduate courses in Acquisition Management and Corporate Entrepreneurship. He has previously taught at the Air Force Institute of Technology (AFIT), where he created the Air Force Strategic Purchasing Graduate Degree Program and served as the director of that program for two years. He has instructed at the University of Dayton, the University of Maryland (University College), and American University (Washington, DC) and at visiting seminars at American University in Cairo (Egypt), and Instituto de Empresas in Madrid (Spain). He has conducted a wide range of acquisition training courses for international military procurement officials throughout the world, including courses in the Czech Republic, Egypt, and Latvia.

Dr. Reed retired from Active Military Service in 2008 after 21 years in the Air Force. He began his Air Force career as a Missile Launch Officer in Minot, ND. He transitioned into the contracting field in 1993 through the Education with Industry Program, working in the F-22 Fighter Engine Development Office at Pratt and Whitney Aircraft Engines in West Palm Beach, FL. He's held various assignments in contracting, including the C-17 Cargo Aircraft and Fighter Engine Systems Program Offices. He served as the director of Joint Contracting Command-North, Kirkuk, Iraq. He also served at the Pentagon as Deputy Chief, Procurement Transformation Division, Headquarters Air Force, where he was responsible for implementing strategic sourcing and commodity councils for the DoD and AF. In his final assignment as Commander, 325th Contracting Squadron, Tyndall AFB, FL, he was responsible for \$500 million in annual purchases in support of F-15 & F-22 fighter aircraft, and AWACS flight training. Dr. Reed earned a PhD in Strategic Management and Entrepreneurship from the University of Colorado and is a Certified Purchasing Manager with the Institute of Supply Management.

Dr. Reed has extensive experience advising organizations in federal government with Censeo Consulting Group. At Censeo, he assisted federal agencies such as Air Force, Navy, Army, Housing and Urban Development, Veterans Affairs and the Executive Office of the President with the creation of sourcing strategies, and provided the training necessary to implement those plans. He is currently the principal director of Beyond Optimal Strategic Solutions in Washington, DC.

He is the author of over 40 articles, book chapters, and conference presentations. His research interests are strategic sourcing, entrepreneurial thinking, acquisition workforce development, and proposal process optimization. He can be reached at tsreed@nps.edu.

Abstract

The increasing pace of change in the acquisition workforce, coupled with a new emphasis on contracting accessions, has increased the interest in the models utilized by the DoD to 1) measure contracting workload, and 2) assign adequate resources to effectively manage the workload with an acceptable level of risk.

Numerous acquisition studies and commissions have cited personnel management as one of the most critical factors contributing to the success or failure of buying organizations. Strategic human capital management and DoD Contract Management have been on the Government Accountability Office (GAO) High-Risk List for the last several years. Actions to understand the nature and dynamics of the acquisition workforce are a first step toward the development and execution of an integrated strategic human capital management plan. Moreover, joint basing and BRAC requirements to merge and



consolidate some contracting offices increases the importance of moving toward a workforce model that is applicable in the joint environment.

This paper identifies the key elements of contracting workforce staffing models in various DoD services. Further, it investigates the rationale and assumptions utilized to develop the models. The validity and applicability of the rationale and assumptions to the current acquisition environment are discussed. The disparity of results and recommendations derived from each of the models, given identical scenarios and inputs, is also identified and discussed.

This research investigates each of the services' workload and resource assessment methodologies in the operational, systems, and contingency environments. Comparisons and contrasts of the various methodologies are discussed.

The research goal is to identify potential opportunities whereby the existing methodologies can be enhanced to reflect the impact of emerging technology on the models in use, ensure the work being performed at various stages within the contract process, and ensure that the differing levels of complexity of the contracting workload are captured in measurement systems.

Keywords: Contracting, workforce, model, human capital



Compensation, Culture and Contracts: The Realities of the DoD's Blended Workforce

Kathy Loudin—Kathlyn Loudin has served on the Defense Acquisition University faculty since 2008. Previously, she led an acclaimed group of cost engineers at the Naval Surface Warfare Center, supporting Navy, Marine Corps, and other DoD programs. Loudin has acquisition experience within both the DoD and industry, holds an MPA, and is a PhD candidate at Virginia Tech. She has written for Contract Management, Defense Acquisition Review Journal, and Defense AT&L, and has presented at the Acquisition Research Symposium (2008), the American Society for Naval Engineers (2008), the Navy Cost Analysis Symposium (2009), and the DoD Cost Analysis Symposium (2010).

Abstract

The Obama Administration's March 2009 mandate that the federal government rebuild its organic acquisition workforce, coupled with the recent repeal of the National Security Personnel System (NSPS), opens the curtain for a fresh look at Department of Defense (DoD) human capital management practices, particularly for employees in critical acquisition positions. Common perceptions hold that the DoD, given its relatively weak ability to provide economic rewards, sacrifices much of its best talent to private-sector employers. Driven in large part by this revolving door phenomenon, the DoD's acquisition workforce now consists of a rich blend of military, civilian, and contractor personnel, who deliver many basic acquisition competencies.

This study synthesizes three quantitative analyses of compensation packages available to military, civilian and contractor personnel, with qualitative research on the less-tangible incentives offered in each sector, to explore the DoD's competitive position in the recruitment and retention of high-caliber acquisition professionals. It finds that, although high-performing DoD civilians are at slight financial disadvantage, they can be motivated to stay in organizations in which a positive, mission-focused ethos prevails.

Background: The Pay-for-Performance Debate

Several years ago, the *National Defense Authorization Act (NDAA)* for fiscal year (FY) 2004 amended Title 5 of the *US Code*, authorizing the Secretary of Defense to implement a new personnel system, known as National Security Personnel System (NSPS), to replace the long-standing General Schedule (GS) system. Heralded as a human capital management system that could boost organizational effectiveness by paying employees for exemplary performance, NSPS was gradually deployed throughout most of the DoD. Six years later, just as the later-adopting DoD organizations were pondering the results of their first NSPS review cycles, the FY 2010 *NDAA* further amended Title 5, effectively rolling back NSPS⁷³ and instituting an improved version of the General Schedule (GS) system, one that will eventually feature management flexibilities and workforce incentives crafted to attract and retain talented employees.

⁷³ The *National Defense Authorization Act for 2010* (US Congress, 2010, HR 2647, Sect. 1113, pp. 309-315) calls for the cessation of NSPS no later than January 1, 2012. This legislation is available at http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=111_cong_bills&docid=f:h2647enr.txt.pdf



Mere months before the repeal of NSPS, in March 2009, President Obama published a memorandum⁷⁴ succinctly addressing the shortcomings of the government's acquisition system and directing the Office of Management and Budget (OMB) to work with Executive Agency and Department Heads to ensure both the capacity and the ability of their workforces to appropriately negotiate, manage and oversee acquisition programs. Within weeks, the DoD had constructed an action plan. This plan called for the addition of some 20,000 acquisition positions by 2015. Of this figure, approximately 11,000 would be converted from contractor positions to government positions; the other 9,000 would be new government positions (Hedgpath, 2009).

Against this exciting but perplexing backdrop, the DoD must balance dramatic demands for experienced, talented personnel, while at the same time reinvigorating the GS system. Some would argue that those two goals are in direct conflict: Many have pointed to the GS system as rewarding longevity (or time in service), rather than retaining the truly high performers. In fact, the main indictment of the GS system was that it failed to motivate strong performance and that it levied few consequences for poor performance. Given that public organizations strive to treat employees equitably, agencies cannot reward stellar performance much more handsomely than they reward mediocre performance. According to James (2002), the GS system has traditionally supported internal equity, but not external (i.e., market-based) equity. Other Office of Personnel Management (OPM) studies have found that 75% of pay increases have been unrelated to performance (Asch, 2005). Administration of pay-for-performance systems is not much easier for private-sector organizations, however. Baker, Jensen, and Murphy (1988) found that the highest-rated employees were only paid a few percentage points more than the lowest-rated employees. Lazear (2000) found a positive correlation only between compensation and time in service and/or hours worked—both objective, not performance-based, metrics.

Recommendations for more effective personnel systems are plentiful: Jamieson and O'Mara (1991) cited flexible rewards, incentive pay tied to both individual achievement and company profits, cash awards for patents and intrapreneurship, stock options, merchandise and travel incentives. Branham (2001) suggested retention bonuses, project bonuses, selective stock options, and higher pay for hard-to-fill positions. However, while the pursuit of more money is a socially acceptable reason for leaving an organization, Branham (2005) found that fewer than 12% of employees actually left for financial reasons. Gellerman (1992) nodded at money as an inducement, but stressed its general inefficiency as a motivator. Paraphrasing Maslowe ("Man does not live by bread alone, except when there is too little bread"), he argued that base pay is essential, but incentive pay motivates the extraordinary.

The DoD's current compensation conundrum seems custom-made for analysis via the principal-agent framework, which emanated from economic theory. In basic terms, principal-agent theory posits that the employee or contractor (the "agent") is rewarded for acting in ways consistent with the best interests of the employer (the "principal"). If the agent's needs are fulfilled, and the principal's expectations are met, then the relationship will endure (Miller & Whitford, 2007). Miller and Whitford also identified a problem dubbed "the

⁷⁴ President Obama's Memorandum on Government Contracting (March 4, 2009) is available at http://www.whitehouse.gov/the_press_office/Memorandum-for-the-Heads-of-Executive-Departments-and-Agencies-Subject-Government/



principal's moral hazard constraint," in which bonuses large enough to produce desired behaviors were cost-prohibitive for the principal. Consistent with much work in this area, I will focus on striking optimal contract(s) among players. Rather than advocate an "either-or" arrangement—i.e., either a behavior-based (base pay) or an outcome-based (incentive pay) contract—I acknowledge the complexity of contracts, both in the real-world sense and in the metaphorical sense. Contracts fall onto a continuum of arrangements, appropriately balancing risks and rewards over time. This lends itself to multiple theoretical lenses for analysis. For instance, studies contemplating risks and rewards associated with (fixed) salary versus (variable) commission packages (Eisenhardt, 1985; 1988; Conlon & Parks, 1988) have augmented principal-agent theory with institutional theory, which addresses the whole hierarchy of human needs. Duncan (2001) leveraged agency theory, alongside equity theory and reinforcement theory, to argue for broad-based incentive stock option plans, because they uphold ownership as a strong source of motivation. Likewise, Blair and Kruse (1999) heralded the rise of employees as an important shareholder group over the past 20 years, embracing employee stock ownership plans (ESOPs) and defined-contribution pension plans, such as 401(k) plans. Wilson (1994) advocated the sharing of rewards, as well, tied to collaboration and teamwork. In addition to the tangible rewards outlined above, compensation structures can also feature intangible or aspirational rewards. Promotional opportunities and career-path options, for example, are often offered along with fixed and variable pay.

Indeed, to create enduring work arrangements, one must not draw solely from economics (a la Hirsch, Michaels & Friedman, 1987). Social relationships are important, as well. Throughout this paper, I will explore embeddedness theory (Granovetter, 1985), which supports the notion of letting managers create a positive climate through pleasing others and generally doing the right thing, thereby serving as role models for the suppression of "force and fraud." The workplace can provide both economic and social incentives for trustworthiness. Such acculturation is an ongoing process, created and calibrated through action in interpersonal networks.

The Blended Workforce

Bringing ever-more complexity to the stage, the phenomenon of the multi-sector, blended workforce has emerged. Contractors, civilians and military personnel are now found working shoulder to shoulder in pursuit of common goals. In a recent study by the Government Accountability Office (2009) of 66 large program offices throughout the DoD, some 37% of their acquisition workforce members were support contractors. Within the Missile Defense Agency and joint program offices, the percentage of contractors was higher: 49% and 47%, respectively. Certainly, from an everyday, operational standpoint, the delineations between "public" and "private" employment have blurred. To understand the nuances of managing in today's complicated blend of mixed allegiances, I decided to conduct a two-part study, balancing the quantitative with the qualitative, and meshing economic with the sociological theory.

Review of the Literature

For decades, researchers—most notably those from the Bureau of Labor Statistics (BLS), charged with Federal Pay Comparability studies—have attempted to compare public and private-sector compensation. Results have been mixed; critiques of various methodologies have abounded. Many have portrayed comparisons of public and private-sector employees as unfair, given the vastly different missions at issue. Bozeman (1987)



argued that because politics infiltrate nearly all organizational behaviors and processes, they are all “public” to some extent. While government and industry differ in their goals, they employ similar mechanisms (a generic set of management functions) in pursuit of those objectives. Allison (1980) concurred that management processes were basically alike, but that the importance of the means (i.e., functions) was eclipsed by the ends (i.e., different missions). Hinting at the “revolving door” phenomenon, Allison cited several high-profile executives⁷⁵ who had performed both public- and private-sector jobs, sharing unanimous sentiment that public management was more difficult.

The difficulties of public management notwithstanding, many practitioners bemoan the relatively low compensation associated with public-sector work. The salary, *per se*, may not be the problem. In fact, Borjas (2002) found that male federal workers earned more than private-sector males with similar experience. The problem is the tight distribution of earnings among public-sector workers: Civil servants performing the roles of greatest responsibility do not earn substantially more than those in less-critical roles. Borjas (2002), Gibbs (2001), and Katz and Krueger (1991) all argued that the compressed distribution of earnings among those in the federal government (relative to the broader possibilities in the private sector) will likely hinder the government’s future ability to recruit and retain highly talented personnel. As many have pointed out, the best employees will always be underpaid; the mediocre will always be overpaid.

Quantitative Analysis: The Surveys

To better assess the DoD’s competitive position with respect to compensation, I collected recent salary data from three professional associations: The National Contract Management Association (NCMA), the Society for Cost Estimating and Analysis (SCEA), and the Project Management Institute (PMI). Members of these three organizations fall into the same general labor categories that commonly make up the DoD acquisition workforce. To normalize the salary and bonus data, I used Consumer Price Indices (CPI) from the BLS website.⁷⁶ Although slight inflation is typically the norm, the purchasing power of \$1 actually increased marginally between 2008 and 2009, so the 2008 salary data are appropriately deflated.

NCMA Survey 2008

The NCMA produced a Salary Survey in 2008, and reported findings based upon usable responses from some 3,543 contracting professionals. Of this sample, approximately 56% were contractors to the federal government, 3% were employed by professional services firms, and 23% were federal employees. Nearly half (47%) worked for very large organizations with annual revenues or budgets exceeding \$501 million.

Of the NCMA sample, 86% had attained at least an undergraduate degree; 45% had earned a graduate degree as well. 42% indicated they currently held professional certifications, such as DAWIA Levels I through III (26%), Certified Professional Contracts Manager (11%), and Certified Federal Contracts Manager (5%). Approximately half reported

⁷⁵ Allison named George Shultz, Donald Rumsfeld, Michael Blumenthal, Roy Ash, Lyman Hamilton, George Romney.

⁷⁶ The Bureau of Labor Statistics inflation calculator is available at http://www.bls.gov/data/inflation_calculator.htm



holding some level of security clearance. More than two-thirds of the respondents resided on the East Coast or the West Coast, typically higher-cost areas. The top cities in terms of median reported salary were San Jose, San Francisco, Los Angeles, and Washington, DC.

As might be expected, salaries correlated positively with age, years of experience, educational level, clearances, professional certifications, and military experience. Interestingly, respondents from the very smallest companies (annual revenues or budget under \$1 million) and from the very largest organizations (over \$501 million) earned the highest median salaries. Of the sample, the typical (median-salaried) male reported a salary \$18,000 higher than the typical female's salary. This difference is likely attributable to differences in experience, education, organization size, and level of responsibility.

Of all of the independent variables, the NCMA researchers identified position as having the strongest relationship to salary. Specific positions reported most frequently were contract manager, supervisor, or director (31% total), and contract administrator or contract specialist (28%). 8% were contracting officers. No other job title was indicated by more than 4% of respondents. Lacking the complete NCMA dataset, the researcher made informed judgments on how titles mapped to the general job functions, reflected in Table 1.

Table 1. Salary by Job Function
(NCMA Salary Survey, 2008)

Job Function	Median Salary	Normalized to TY09\$	% in Function
Executive	\$ 152,000	\$ 151,459	4% \$ 6,058
Attorney	\$ 140,000	\$ 139,502	1% \$ 1,395
Manager	\$ 115,000	\$ 114,591	18% \$ 20,626
Consultant	\$ 112,000	\$ 111,602	3% \$ 3,348
Contract Mgr	\$ 108,000	\$ 107,616	15% \$ 16,142
Subcontract Mgr	\$ 104,000	\$ 103,630	15% \$ 15,544
Supervisor	\$ 99,800	\$ 99,445	16% \$ 15,911
Staff (Contract Admin.)	\$ 76,700	\$ 76,427	28% \$ 21,400
Weighted Average, All Functions (TY09\$)			\$ 100,425

In comparing median salaries of contractors to those of federal employees, pronounced differences were not found. Contractors reported salaries approximately 2% higher than those of federal government employees. Consultants (i.e., employees of professional services firms) earned considerably more; however, lacking complete data from the NCMA study, it was not clear whether these consultants were working with government or commercial clients.

Table 2. Salary by Type of Employer
(NCMA Salary Survey, 2008)

Employer	Percentage	Median Salary	Normalized to TY09\$
Contractor to Government	56%	\$ 92,000	\$ 91,673 \$ 62,606
Federal Government	23%	\$ 90,000	\$ 89,680 \$ 25,154
Consultant	3%	\$ 100,000	\$ 99,644 \$ 3,646
Weighted Average, Certain Employers (TY09\$)			\$ 91,405

Analyses of other elements of compensation packages revealed compelling differences. While nearly all respondents (96%) received some package of paid absences, 92% were offered healthcare assistance, 90% were entitled to participate in a 401K or similar plan, 88% were eligible for life insurance, 85% were provided with dental care



assistance, 75% qualified for vision care assistance, and 70% were afforded tuition assistance; other valuable benefits were bestowed upon smaller segments of the sample. Only 35% were entitled to a pension plan, 30% (contractors, consultants, and commercial-business employees only) were offered ESOPs, and 64% were eligible for bonuses. For the respondents eligible to receive bonuses, the median bonus was \$3,500 (TY08\$). 24% reported a bonus of \$10,000 or more, 9% indicated a bonus of less than \$1,000, and 3% indicated none. Of respondents identified as executives (roughly 4% of the sample), the median bonus was \$25,000 (TY08\$). Bonuses by percentile are shown in Table 3.

Table 3. Bonuses for Contracting Professionals
(NCMA Salary Survey, 2008)

Bonuses (normalized to TY09\$)	
10th percentile	\$ 794
25th percentile	\$ 1,589
50th percentile	\$ 3,475
75th percentile	\$ 9,631
90th percentile	\$ 19,858
Executives	\$ 24,822

SCEA Survey 2005

The most recent salary survey from SCEA was conducted in 2005, and reflected usable responses from 405 professionals. Of this sample, approximately 78% were male and 22% were female. Geographical data were expressed in terms of SCEA chapter affiliations; because 25% of respondents did not identify a specific chapter, the prevalence of respondents' geographic locations could only be roughly estimated. Approximately 42% of SCEA respondents were linked to either East Coast or West Coast chapters (typically in higher-cost areas). Of the SCEA sample, 67% worked for private-sector companies and 30% were government employees. Of the government employees, 89% were civilians and 11% were active-duty military personnel; of the military personnel, 92% were Air Force and 8% were Navy. For the private-sector respondents, no breakouts were provided on type or size of company. Instead, breakouts of primary end products were given. The greatest number of participants (30%) worked aircraft, missile and spacecraft production. The next-largest category was research and consulting (24%), followed by electronics (11%), and intelligence/reconnaissance (7%).

SCEA respondents were highly educated: 97% had attained an undergraduate degree; 70% held a graduate degree as well. 36% indicated that they had earned SCEA's professional certification. Again, salaries correlated positively with age, years of experience, and educational level. Of the sample, males generally reported significantly higher salaries than females, even within similar geographic areas and job functions. For example, among the 65% of all respondents identified as cost estimators, males earned 16.5% more than females. Disparities such as this are likely tied to differences in experience and responsibility levels: The typical (i.e., median) male respondent reported 19 years of experience, while the typical female reported 15 years. The males were slightly more likely to shoulder supervisory responsibilities as well.

Although the vast majority of SCEA participants were cost estimators, salary data were reported by several other types of professionals. Program managers and financial managers comprised 10% and 11% of the sample, respectively. Contracting and Earned



Value Management (EVM) professionals reported the highest salaries, but collectively represented only 6% of the SCEA sample, as shown in Table 4.

Table 4. Salary by Job Function
(SCEA National Survey Results, 2005)

Job Function	Median Salary	Normalized to TY09\$	% in Function	% of Total Reporting Salaries	
Earned Value Management	\$ 114,000	\$ 125,229	3.00%	0.0326087	\$ 4,084
Contracting	\$ 108,000	\$ 118,638	3.00%	0.0326087	\$ 3,869
Program Management	\$ 100,000	\$ 109,850	10.00%	0.10869565	\$ 11,940
Financial Management	\$ 94,450	\$ 103,753	11.00%	0.11956522	\$ 12,405
Cost Estimating	\$ 92,030	\$ 101,095	65.00%	0.70652174	\$ 71,426
Accounting & Other	not reported	not reported	0.00%	-	\$ -
Weighted Average, All Functions (TY09\$)					\$103,723

Overall, as was the case with the NCMA survey, the median salaries of SCEA's private-sector employees were closely aligned with those of government employees, averaging about \$107,000 (TY09\$). Interestingly, though, with increasing levels of experience, salaries of public-sector professionals greatly surpassed those of private-sector employees, as shown in Table 5.

Table 5. Salary by Type of Employer and Years of Experience
(SCEA National Survey Results, 2005)

Employer & Experience Level	Percentage	<10 years	10 to 19 years	20 to 29 years	≥ 30 years	Median Experience Level = 17 years	Median Salary Normalized to TY09\$
Business	67%	\$ 71,500	\$ 98,000	\$ 98,000	\$ 95,000	\$ 98,000	\$ 107,653
Government	30%	\$ 75,000	\$ 97,000	\$ 106,000	\$ 108,000	\$ 97,000	\$ 106,554

Aside from base salary, the SCEA survey did not address any other aspects of compensation, such as bonuses and fringe benefits.

PMI Survey 2007

The Project Management Institute (PMI) released a salary survey in 2007, reflecting usable responses from some 1,143 professionals from the United States. Of the PMI sample, approximately 65% were male and 35% were female. Geographical data were not collected by specific location, but by type of locale. Nearly two-thirds of respondents (65%) worked in large cities, while roughly one-third (31%) worked in small to medium-sized cities; the remaining 4% worked in rural areas. Of the PMI sample, 83% worked for private-sector companies; 9.5% were government employees; 7.5% were consultants. Nearly half (44%) of the respondents worked for very large organizations (i.e., more than 10,000 employees).

Of the PMI sample, 88% had attained at least an undergraduate degree; 44% possessed a graduate degree as well. 64% indicated that they held PMI's professional certification. As expected, salaries correlated positively with age, years of experience, level of education and certification. PMI also mapped salaries to project size, project team size, and project budget, both of which positively correlated with salary. Of the sample (65% male, 35% female), salaries for males were roughly 10% higher than those for females. PMI respondents, not surprisingly, were predominantly project and program managers. To ensure definitional consistency across the country, PMI detailed the functions for each title within the management hierarchy, and directed respondents to select the one most closely



approximating their normal job duties, rather than reporting on a company-specific (or contract-specific) title currently held. For example, PMI (2007) distinguishes program manager from a project manager III as follows: A program manager coordinates multiple, interrelated projects in pursuit of a common operational objectives, while a top-level project manager oversees high-priority projects, involving extensive functional integration and considerable resources.

Table 6. Salary by Standardized Job Title
(PMI Salary Survey, 2007)

Job Title	Median Salary	Normalized to TY09\$	% in Role	
Chief Executive Officer	\$ 130,000	\$ 134,511	1%	1,295
Chief Information Officer	\$ 125,000	\$ 129,338	1%	1,471
Director of PMO	\$ 120,000	\$ 124,164	4%	5,323
Portfolio Manager	\$ 111,065	\$ 114,919	6%	6,535
Program Manager	\$ 105,500	\$ 109,161	22%	23,685
Consultant (Internal or External)	\$ 104,384	\$ 108,006	6%	6,709
Functional Manager	\$ 100,000	\$ 103,470	5%	4,888
Project Manager III	\$ 94,000	\$ 97,262	28%	26,975
Project Manager II	\$ 86,100	\$ 89,088	15%	13,172
Specialist (Scheduler, Cost Analyst)	\$ 84,500	\$ 87,432	4%	3,672
Project Manager I	\$ 82,750	\$ 85,622	8%	7,041
Weighted Average, All Jobs				\$100,766

Consistent with the preceding two surveys, PMI's median-base salaries differed only slightly across sectors. PMI collected salary data by industrial category; the categories typical of DoD contractors (information technology, aerospace, engineering, manufacturing, and telecommunications) are presented in Table 7. As was the case with the NCMA survey results, consulting firms appeared to pay most generously; however, it was not possible to tell whether the consultants were doing business with the government or with other private companies.

Table 7. Salary by Type of Employer
(PMI Salary Survey, 2007)

Types of Employer	Percentage of Sample of Interest	Salary	Normalized to TY09\$	
Consulting	8%	\$111,500	\$115,369	\$ 15,643
Aerospace	3%	\$104,500	\$108,126	\$ 5,498
Telecommunications	6%	\$101,000	\$104,505	\$ 10,628
Engineering	5%	\$98,975	\$102,410	\$ 8,679
Government	10%	\$98,000	\$101,401	\$ 17,187
Manufacturing	5%	\$97,000	\$100,366	\$ 8,506
Information Technology	22%	\$93,000	\$96,227	\$ 35,881
Weighted Average, Certain Employers (TY09\$)				\$ 102,021

Only on the aggregate level did the PMI survey address bonuses and other incentive pay. Since PMI defined "total compensation" as salary plus bonuses, Table 8 displays the delta between total compensation and base salary for all United States respondents; these figures reflect all sectors and all job titles. Because the mean figure can be distorted by



outlying data (e.g., extremely high executive bonuses), the median (i.e., 50th percentile) figure will be used in subsequent comparisons.

Table 8. Bonuses for Project Management Professionals
(PMI Salary Survey, 2007)

Bonuses (normalized to TY09\$)		
25th percentile	\$	4,656
50th percentile	\$	7,295
75th percentile	\$	13,192
Mean	\$	11,035

Synthesis of Three Surveys

The PMI, SCEA and NCMA surveys were selected because they focused upon the exact types of employees desired for this study (i.e., program managers, cost estimators, and contract managers). As detailed in the preceding sections, the preponderance of participants in each of the respective surveys fell into those named categories. Summarized in Table 9 are the median base salaries (TY09\$), as reported for government organizations, consulting agencies, and the types of companies likely to be DoD contractors (e.g., aerospace, information technology, engineering, telecommunications, and manufacturing firms). Since only certain types of employers were deemed of interest within each sample, the percentage columns do not always add up to 100.

Table 9. Three-Survey Comparison of Base Salaries and Bonuses for Project Managers, Contracting Professionals, and Cost Estimators (TY09\$)
(PMI, 2007; SCEA, 2005; NCMA, 2008)

Type of Employer	PMI Survey		NCMA Survey		SCEA Survey	
	% of Sample	Med Salary	% of Sample	Med Salary	% of Sample	Med Salary
Contractor	41%	\$96,229	56%	\$91,673	67%	\$107,653
Government	10%	\$101,401	23%	\$89,680	30%	\$106,554
Consulting Firm	8%	\$115,369	3%	\$99,644	0%	
Median Bonus		\$7,295		\$3,475		n/a

Clearly, these figures cast doubt on the common perception that government employees are poorly compensated relative to their contractor counterparts. However, it should be noted that, in each of the three samples, the percentage of contractor employees represented far exceeds that of government employees. It could be inferred that fewer government employees belong to professional organizations, or that few feel compelled to report their salary data, since federal civilian and military pay data are publicly accessible.

Moreover, more analysis is necessary to pinpoint compensation data on the *expertise levels* needed within each labor category to manage major defense acquisition programs (MDAPs). To provide a more focused comparison of the compensation of a DoD contractor to DoD government (both military and civilian) employees, at the levels customarily involved in managing MDAPs, the researcher constructed the chart in Table 10.

Table 10. Cross-Sector Comparison of Total Compensation for Employees Supporting Major Defense Acquisition Programs (TY09\$)

	Military	Civilian	Industry	Military	Civilian	Industry	Military	Civilian	Industry
	<i>Project</i>			<i>Program</i>			<i>Exec</i>		
	<i>O-4</i>	<i>GS-13</i>	<i>Manager</i>	<i>O-5</i>	<i>GS-14</i>	<i>Manager</i>	<i>O-6</i>	<i>GS-15</i>	<i>Director</i>
<i>Basic Pay</i>	78,408	96,469	94,048	82,248	113,998	106,489	87,480	134,094	131,219
<i>Bonus</i>	0	1,300	4,205	0	1,400	4,761	0	1,500	5,867
<i>Healthcare</i>	12,000	6,900	6,900	12,000	6,900	6,900	12,000	6,900	6,900
<i>Housing</i>	22,884	0	0	24,948	0	0	25,200	0	0
<i>Subsistence</i>	2,676	0	0	2,676	0	0	2,676	0	0
<i>Tax Savings</i>	6,390	0	0	6,906	0	0	6,969	0	0
<i>Retirement</i>	3,920	4,823	3,292	4,112	5,700	3,727	4,374	6,705	4,593
<i>Paid Absences</i>	8,444	12,295	8,114	9,031	14,529	9,187	9,606	17,090	11,321
<i>Total</i>	\$130,803	\$116,964	\$113,267	\$137,809	\$136,827	\$127,337	\$143,931	\$159,584	\$155,306

This chart captures the mix of acquisition personnel who typically support major DoD programs. Given the complexities inherent to the MDAP environment, it was assumed that contractor capabilities would be *at least* equivalent to a project manager II or III (in PMI terms); the interpolation of that salary aligns closely with the weighted average (\$94,048 in TY09\$) of median contractor salaries reported by the NCMA, SCEA, and PMI. Similarly, the median bonus mapped to an employee at that level reflects the weighted average (\$4,205) of bonuses reported by the NCMA and PMI. Within the MDAP framework, the second tier of contractor employee is equates to a program manager (in PMI terms). Since the median program manager salary was 13.23% higher than the project manager II/III salary, commensurate adjustments were made to the salary and bonus at second level. Similarly, the top tier of contractor employee aligns with an executive or a director (depending on the company). Since that salary was typically 23.22% higher than a program manager salary, appropriate adjustments were made to the contractor salary and bonus at the top level. Support for this methodology was obtained through a series of personal interviews with both contractors and civilians affiliated with MDAPs in early 2010; more details are provided in forthcoming sections.

For civilian salaries, all figures in Table 10 reflect base salaries at the midpoint of the grade level under consideration, augmented with the national average of 20.54% in locality pay.⁷⁷ For the military pay computations, base salaries for officers with 14 years of experience were used.⁷⁸ Military housing allowances⁷⁹ for Minneapolis were chosen due to that metropolitan area's similar cost of living; it most closely approximates the national average with 20.36% in locality pay. The basic allowance for subsistence for all military officers is \$223 per month.⁸⁰ Because DoD payments of housing allowances and subsistence are not taxable, military personnel save approximately 25% of that combined amount on income taxes each year, relative to their civilian and contractor counterparts.

⁷⁷ Locality pay data (2009) are available at <http://www.opm.gov/oca/09tables/indexGS.asp>

⁷⁸ Military pay tables (2009) are available at <http://www.dfas.mil/militarypay/militarypaytables/2009MilitaryPayTables.pdf>

⁷⁹ Housing allowances (2009) are available at <http://www.defensetravel.dod.mil/perdiem/bah.html>

⁸⁰ Basic allowance for subsistence is available at <http://militarypay.defense.gov/pay/bas/>

Valuation of non-monetary compensation involved making some assumptions. For example, healthcare is provided free of charge to military personnel and their dependents. Assuming the total premium cost plus medical care would average \$1,000 per month per family, the annual benefit to military families was assessed at \$12,000. While civilians and contractors generally receive assistance on healthcare premiums as well, the researcher assumed an out-of-pocket cost of \$2,600 per year for insurance premiums and \$2,500 per year for medical costs. Accordingly, the valuation of the contractor and civilian-employee benefit was \$6,900.

For matters concerning retirement benefits, there are endless variations on their valuation and administration. For the purposes of this analysis, however, a simple approach was used. For civilians and military personnel, the full potential agency contribution, equating to 5% of annual salary, was used. For industry employees, a notional company contribution of 3.5% was used.

Crisp calculations of the value of paid absences were stymied by the methods of accounting for such benefits. In the military model (a 24-hour, 7-day-per-week operational context), personnel are granted approximately 30 days off per year, but additional leave is sometimes authorized. On the other hand, additional duty days are often required. In this comparison, 25 days was chosen as representative of the leave utilized by military personnel. Active-duty personnel are also entitled to unlimited sick leave; for this comparison, three days of sick leave per year was assumed. For each of the three civilian grades, annual leave was valued at six hours per pay period (the level afforded to employees with fewer than 15 years of federal service).⁸¹ For the civilian grades, it was assumed that the average employee takes three days of sick leave annually. Finally, federal holidays added 10 days per year to the paid-absence total for civilian and military personnel. For paid absences in the private sector, a factor of 22 days per year (combined total of family/personal, sick leave, and federal holidays) was used.

Qualitative Analysis: The Interviews

To check the realism of these assumptions, while collecting qualitative data on the non-monetary factors that help retain high-quality personnel, I embarked upon more than 30 in-depth interviews with both public- and private-sector managers during the months of February and March 2010. My research study was publicized via LinkedIn.com, a career-oriented networking site,⁸² to several DoD-oriented groups, as well as through word of mouth to other colleagues. Targeted participants were managers with a minimum of 15 years of experience, working for at least two of the three types of employers (military, DoD civilian, or contractor). My interview questions were designed to capture both sides of the principal-agent relationship. After all, managers are familiar with both sides of that equation: They serve as principals on behalf of their organizations, but they perform as agents themselves.

Interviewees were given read-ahead material, so that the interviews could be conducted within a 45-minute timeframe, either face-to-face or via telephone. Given time constraints, I had to solicit input via e-mail in a few cases. Lacking a two-way mechanism

⁸¹ While most senior civilians have more than 15 years of service, it is probable that their program-office duties preclude the full use and enjoyment of leave benefits; accordingly, the figure used for comparison purposes was 6 hours per biweekly pay period (or 19.5 days per year) of annual leave.

⁸² See, for example, http://www.linkedin.com/groups?gid=67539&trk=myg_ugrp_ovr



for clarification and feedback in those exceptional cases, the questionnaires that were delivered via e-mail were reviewed thoroughly for consistency and realism. A few responses had to be discarded for reasons of inconsistency. While the interview questions were standard, they could be altered in order to fully explore the interviewees' knowledge areas. In very few situations, the focus remained on financial compensation; in most cases, the range of topics encompassed a host of other intangible inducements, including organizational culture.

The interviews took place in a variety of venues—in the cushioned alcoves of briskly percolating cafes, under the harsh lighting of 1940s-era government facilities, in modern high-rise offices, in the lobbies of posh hotels, behind the closed doors of unused conference rooms, and in quiet corners of company cafeterias. Due to my limited travel budget, however, just as many interviews were conducted via telephone—from my desk in northern Virginia, from telework centers, and from my kitchen table.

Compensation, I affirmed at the beginning of each interview, is a term that extends well beyond figures on a paycheck. Compensation can be understood as the complete package of rewards for work performed in support of an organization. The total package may include ESOPs, profit-sharing, bonuses, tuition assistance, rewards for improving credentials, subsidized childcare, transportation or food service, healthcare, retirement contributions, life insurance, housing, subsistence, travel, and relocation assistance, as well as flexible working arrangements (hours, location, job-sharing), home-office equipment, car allowances, and reinforcements such as event tickets, gift cards, and free coffee, soda and snacks (reminders that the organization is constantly paying you back for your hard work).

“Industry Pays More” ... Or Does It?

In the Washington metropolitan area, one manager stated, “Industry jobs are \$155K and up. With ‘tickets,’ you can command another \$20K. Since the DoD salary range is now around \$30K to \$155K, it simply cannot compete for people *who are after high salaries.*” Another manager corroborated this: “As a person with ‘tickets’ and certifications, I make \$175K per year.” One manager with visibility into salaries (both CONUS and OCONUS) stated that program manager salaries range from \$120K and \$150K, with clearances and professional certifications potentially adding \$5K to \$20K to the base salary. Yet another interviewee acknowledged that “top-level program managers make \$160-\$200K. Of course, higher-level positions pay much more, about \$300K and up.” A civilian manager with cognizance over contractors observed that “nearly all senior-level industry folks, with or without managerial duties, make more than \$100K here. The program managers I work with are pulling down \$160K to \$180K in base pay.”

While most interviewees applauded the pay flexibilities offered by industry, not all were able to secure high salaries initially. “When I retired (as an O-5) and interviewed with a DoD organization that used paybanding, they equated my rank to a GS-12 or 13. I made a sacrifice in terms of earning power to work within DoD, but I did not stay there.” Similarly, a former Army officer recalled questionable salary advice while transitioning. “My counselor recommended scaling back my pay expectations, but I knew that the only way to something close to the salary I wanted was to ask for it!” Other retirees mentioned trade-offs: “Since I didn’t need healthcare insurance, I traded those for a higher salary.” “When I moved to industry, there was more flexibility with salary, but I couldn’t negotiate more leave. The company was very concerned about overhead rates, but high direct rates seemed to be okay.”



“Non-taxable benefits are a big draw to me now. When first I left the military, I was shocked to see a large portion of my paycheck being taxed away.” To attract contractors to support military operations in both hostile and non-hostile environments, one manager indicated that overseas employees were entitled to tax savings of \$28,000 from non-taxable compensation.

In industry, according to another manager, compensation is far more subjective. “I’ve seen salaries differing as much as 20% for people doing technically the same job. I’ve seen vice presidents with almost a 100% difference in salary for the same job—at these higher levels, it comes down to this person’s contribution to the overall mission. That’s hard to quantify because a lot of ‘soft’ assets—creativity, leadership and decision-making skills—come into play.” Another manager stressed that salaries for *good people* will be higher in industry than in government. “You get what you pay for. If you have the good fortune to be able to hire the best people, do it! You can accomplish just as much with one great person as other companies can do with two or three average people.”

Why, then, would a highly motivated performer want to work within the DoD? Most of the interviewees pointed to the intrinsic rewards. “When I retired, I first worked as a contractor. However, my leadership capabilities—those I had been refining throughout my entire military career—were just not used in that contractor role.” Other retirees concurred. “It’s very hard, after being a key decision-maker, to just be an advisor, subservient to inexperienced program-office personnel who can easily dismiss my opinion.” Another downside for military retirees joining private companies is that, in order to win business, they are often required to exploit relationships with people still working within the DoD. “The pressures that come with generating business are sometimes not worth the extra pay.”

Bonuses and Awards

Bonuses and awards, given their highly variable nature, are among the most interesting aspects of compensation. One manager related that he had held positions with four different contractors, but had rarely heard of five-figure bonuses. Another manager said that bonuses were relatively rare, but that everyone was eligible. “Our bonuses are based on percentage of Award Fee dollars earned. Our goal is to delight the customer, so our incentive is tied to that.” Other managers revealed that bonuses are not typical: “Only the people who have invested at least 10 years with the company are eligible for bonuses.” “No one gets bonuses unless they’ve been with the company for a very long time.” “In the companies I’ve worked with, the program manager did not get any special bonus.” “Only the most senior program managers get bonuses.”

Other companies employed very different practices with respect to bonuses. “We notify employees at the beginning of their performance-assessment cycle of their eligibility. To get the bonus, there’s a range of compensation and goals/objectives to meet. At the end of the cycle, feedback is provided, and a final payout decision is made.” Another manager affirmed that “everyone is eligible for some type of bonus, based on our sector’s profitability.” Still another stated, “Bonuses are based on project performance. Meeting revenue and profit marks set by upper management constitutes half of the bonus, while the other half is subjectively determined by the direct supervisor.” Another manager was pleasantly surprised: “I haven’t been with the company through the entire year, but I just received \$2,800—after just four months on the job.” Two other managers said, “Annual bonuses are an expectation of all employees.” “Bonuses are a great retention tool, but have to be administered carefully.” One manager uses multifaceted bonus structure, based on



customer satisfaction, business development, collateral duties, and participation in morale-building activities.

On the government side, civilian bonuses depend less on performance than on the size of the organization's pay pools. "Under NSPS," said one manager, "individuals could be rewarded more fully. However, to many of my senior folks, being recognized as a 'five' (top of the scale) is more important than money." Another manager concurred: "Disappointments among my engineers tend not to center on the paltry bonuses, but on not being rated 'outstanding.' With limited bonuses to go around—and personnel rules tying the amount to the rating—not everyone can achieve that top rating." Yet another civilian manager confirmed that the small bonuses are not enough to influence behavior: "We're usually talking about 1% of salary here. The good part is that people know and understand this." Military officers, on the other hand, can obtain special pay for specific job duties, but are not eligible for bonuses. Instead, officers are motivated by the potential of future promotions.

When asked about group bonuses to reward collaborative efforts, only one interviewee had been part of an organization that gave out group bonuses, and they represented a very small percentage of the overall compensation strategy. To encourage cohesiveness and teamwork, many of the managers bestowed non-monetary recognition upon groups, during periodic award ceremonies, through letters of recognition, and within the context of monthly project reviews.

Salary Reviews

When asked about the frequency and the impact of salary reviews, one civilian jokingly feigned confusion over the question. "I do what I do, and by all accounts I do it well. Nothing ever changes my salary—except cost of living adjustments." Another civilian provided clarity: "Because DoD organizations have become fairly flat, many technical experts can only go so far." Others concurred: "For military and civilians, performance reviews exist, but salary reviews really do not."

Within industry, salaries are usually reviewed annually, or "upon negotiation for an increase." Said one manager: "Most companies have an out-of-cycle process for unique situations that require immediate address." Increases of 3% to 15% per year were reported, with the higher increases afforded to lower-salaried personnel. Sometimes, though, pay increases do not keep up with the cost of living. "This year, given economic concerns, our company elected to freeze all salaries over \$100K." Another manager revealed, "Depending on the economy, we get a 3-5% pay bump. But anyone who's been in this business awhile knows that the best way to increase your income is to play the mercenary role: Move over and move up! You don't gain a whole lot in terms of earning potential by staying with the same company." Others agreed: "Sometimes, the only way to boost compensation is to leave your current company for a higher bidder. Accordingly, industry employees often have a better sense of their market value. With that comes a drive to work harder and move forward."

One manager pointed to an unintentional seniority disincentive. "If the average pay increase for 15 years was 2% per year, while salaries for new hires increased at 2.2% per year, eventually you will have new hires making more than the old timers, since prevailing rates were paid to lure new people, but only minor adjustments were made to keep the old." Such problems are solvable, given the flexibilities of contractor systems. One manager touted the opportunity to really look out for people, where appropriate. He has autonomy with salary decisions, but is also accountable for profits, losses, and customer satisfaction.



“One time, an employee bargained for a 30% pay increase. He was working on a cost-plus contract. There’s no way I could ask the government to cover that, so I let him go and get that raise from another company.”

Another manager reported conducting salary reviews after just six months for exceptional new hires, but emphasized that “salary reviews, by themselves, are not good retention tools. Other things seem to matter a lot more.” Clearly, retention of talent is not contingent solely upon economics.

Other Aspects of Compensation: The Work Itself

Fueled by technological advances, workplaces have fundamentally changed over the past decade. Flexible working hours, mobile communication devices, and telework arrangements are largely taken for granted. As such, they no longer hold great motivational power. In 2010, the most frequently cited non-monetary forms of compensation centered upon aspects of *the work itself*. A civilian pointed to travel opportunities. “My projects are all over the country, so I’m rarely stuck in my office for more than a few weeks at a time.” Another nodded to his colleagues. “They are well educated and motivated. We don’t have to worry about back-stabbing from the inside! Sure, there are politics everywhere, but here we don’t tend to have ‘camps’ and ‘cliques,’ as I’ve witnessed when I spend time at my contractors’ facilities.” Other civilians concurred: “It’s definitely not Us versus Us. Friction might fester between Us and Outsiders (e.g., Congress and resource sponsors), but this helps bolster that sense of shared mission. It keeps us together.”

“My military career prepared me to move—had to go in, learn the ropes, and perform wherever I was sent. Two or three years later, I had to pack up and do it all again. To be successful, you needed agility of mind, and willingness to take on more responsibilities and manage risks.” These experiences primed her for senior positions, which she found in the private sector, but has since moved back to the DoD, where she feels she can directly improve the livelihoods of soldiers and sailors. Another retired officer emphasized intrinsic motivators. “After 20 years of leadership acculturation, I gravitate to positions where I am in charge.” Having worked as a support contractor, “I couldn’t stop program offices from making bad decisions. Now I can!” Another retiree who has worked across sectors cited satisfaction in teaming with other strong female civilians, as well as people from diverse socioeconomic backgrounds. “Opportunities for this are far greater within DoD than in the traditional white, male, upper-middleclass network found in industry.”

A civilian manager stated that “many of my senior folks hold patents and doctorate degrees. Accordingly, they command fairly high salaries, but that’s not why they work. They just love what they’re doing!” Similar sentiments came from another civilian manager, for whom money is not a big motivator. “I feel that I’m making a positive difference. Every morning I wake up energized to go to work and do great things!” As highly credentialed professionals move through their careers, money may become less important. A civilian manager mentioned cases in which senior systems engineers took \$50-70K pay cuts to come back to the DoD. “They wanted ‘quality of life’ things—respectful working environments and manageable workloads.”

Certainly, job satisfaction can be found outside of the DoD as well. One industry manager reminisced, “I’ve held jobs where I just loved the work—I was making a difference in soldiers’ lives, because I was right there, giving them the tools they needed.” For another, “in taking the reins of a small company, I was trusted to make decisions without constraints. Somehow, the stockholders knew that I would do the right thing! My goals were simple—to grow the business and protect the employees from external shocks.” A senior civilian



agreed that “too many rules lead to managers becoming robotic—just going through the motions—people need some latitude to be creative problem solvers.”

Managing the Blended Workforce: Different Norms, Different Expectations

When queried about the relative ease of managing civilians versus contractors, most interviewees stated that contractors could be relied upon to deliver results within the timeframe allocated. If managed astutely by a strong Contracting Officer's Representative (COR), the contractor will have firm deliverables, a well-defined scope, and established schedules. This is less common on the government side, where deliverable dates tend to be flexible, budgets are complex, scope is virtually unlimited, and schedules stretch for many years. One civilian offered that contractors “don't see the breadth of our responsibilities. Whereas a contractor might provide great support in a well-defined area, we are all over the place—steering diverse projects and programs—managing enormous amounts of money over long timelines.” An industry manager offered that “DoD civilians face tremendous administrative hurdles in getting things done.”

“There's definitely more motivation for contractors to perform,” said a manager with experience on both sides. “That's due to a higher paycheck, supposedly. But I think there's more to it. There's more focus—and a greater sense of accountability.” This could stem from less job security. “Industry projects are very time sensitive. Failure to make budget or meet deadlines usually translates to dismissal,” said one manager. “If you screw up, you are history,” said another. “We have to take calculated risks in order to make things happen,” another industry manager said. “Out here, you've got to *produce*...and quickly...making decisions and meeting customer demands.” Managers with experience on both sides agreed that there are many culture-driven expectations regarding employees' work habits. “Uncompensated overtime in industry is normal. I log about 10 extra hours per week, but that's just what I do as a professional.” On the civilian side, “There's no expectation that anyone should work more than a 40-hour week. In the working-capital environment, the norm of working no more than 80 hours per two-week pay period was especially pronounced, because labor accounting there had to be a lot more precise.”

A civilian shared a different perspective. “My Blackberry-equipped program managers are almost always on duty. I can send them a question any time of the night, and I'll have an answer within the hour.” Although managers from both sides were quick to note DoD civilians who personified dedication and commitment, the general thinking was that “Civilian workers are a little bit too secure.” “It's hard to fire them—there are no grave consequences when they mess up.” “There's more acceptance of mediocrity.” Disturbingly, one manager pointed to “few incentives for departing DoD experts to share knowledge with the junior folks, who are discouraged when they are not being challenged.” Civilian time-in-service promotions are another irritant: “People can move all the way up to GS-13 without necessarily performing strongly.” Two contractors expressed sadness over the repeal of NSPS. “Government managers really need a ‘stick’ to keep people productive. (They need better ‘carrots,’ too.)”

Broadly speaking, performance expectations for military personnel and industry workers were perceived as being higher than they are for civilians. These differences are driven by job-security issues (on the industry side), and the ability to gain promotions in rank (on the military side). Some observed that “the culture within industry is similar in some ways to active duty. You do whatever it takes to get the job done.” A civilian manager disagreed, citing no significant differences in terms of responsibility or organizational



expectations. “As a program manager, I always had a counterpart on the industry side with similar duties and concerns. We understood one another and worked well together. Whenever our respective organizations had differences, we let the attorneys duke it out and we continued our collaborative engineering!”

“There are no problems with the blended workforce, as long as there’s good leadership.” A civilian manager agreed: “People are people—we should adopt a ‘colorless badge’ ethic.” An industry manager confirmed that “values are embedded in people, regardless of where they work. That’s what the DoD customer is really buying—great people, strong ethics, professional judgment. Of course, there are lines that cannot be crossed. We develop products for decision-makers to use—Independent Cost Estimates, Acquisition Plans, Performance Work Statements, etc.—but we cannot support a program that our company might eventually bid on.”

Hiring, Firing and Everything in Between

“No rational private company would institute a personnel system like ours,” confessed one civilian manager. Most DoD organizations have a fixed number of billets, and cannot initiate *ad hoc* hiring actions. Processes must be followed; attritions are not automatically backfilled. Another manager admitted, “when I worked in the private sector, I appreciated the speed and smoothness with which human resources functions were carried out.”

Several retired officers were repelled by the application process for DoD jobs. “So much documentation, so many narratives on knowledge, skills and abilities (KSAs)—who actually reads this stuff?” When applying for DoD jobs, “you upload documents to a ‘black hole’ of a database, and don’t hear anything for months! I considered DoD when I retired, but I have to put bread on the table. I can’t wait six months for a call.” Radically different stories came from this industry side: “All I did was post my resume on the NCMA website, and within days, I had 20 potential interviews. The first inquiry came within hours!”

“We have lots of flexibility to recruit and retain...without a lot of hassle,” said one manager. With this discretionary power, though, comes perceptions of reduced transparency. Another manager stressed, “We seem to have few standards on years of experience needed, or the value of degrees and certifications. It’s very subjective, and when employees ask what they need to do in order to get promoted, I cannot offer much guidance, except to keep performing!”

“On firing, our rules are obviously looser than DoD’s. As a director, I could fire for cause (e.g., harassment) or for non-performance. I could fire any of my direct reports *on the spot* for serious infractions (e.g., assault).” As chronicled in the preceding section, many managers are concerned over difficulties in terminating non-performing civilians. “Deadwood does exist,” stated one. “Non-performers get shuffled around, but only rarely do we find a better fit that fixes the performance problem.” “Unfortunately, the lack of productivity from poor performers puts more pressure on the civilians who do try to make things work,” a civilian explained. “This can lead to frustration and the exodus of good employees.”

Other examples of inflexible rules emerge when DoD civilians obtain attractive offers from other DoD organizations: The employing organization cannot respond. “It’s hard to keep my best folks from being lured away,” said a senior manager. “I must recruit very strategically, targeting, for instance, talent made available up by a Base Realignment and Closure (BRAC).”



Proactive succession planning is also critical. “We mentor *all* GS-14s and GS-15s,” said a civilian manager. “That way, no one feels singled out or neglected. These are our senior folks—the ones next in line to lead. Whether or not they get that promotion, they’re the ones running departments and divisions now, so they need to stay energized.” Mentoring younger workers is also key. “With Generation Y, the ‘thou shalt stay’ mentality is counterproductive. We need to engage in enterprise-level thinking. Every person needs to find the best environment in which to use their skills. For some, it’s within DoD; for others, it’s with a DoD contractor. By taking this broader view, we can encourage young people to find their own way, without stifling growth and confining them to stovepiped organizations.”

Compensation, Culture and Contracts: Conclusion

This study was conducted in two phases, which sometimes overlapped and required additional iterations. While the first phase involved gathering, normalizing and organizing the quantitative data, the second was aimed at gaining a richer understanding of the data via qualitative interviews with highly experienced managers. My goal was to produce a balanced view of compensation practices, situated within particular organizational cultures, thereby infusing principal-agent theory with embeddedness theory. From a sheer economic perspective, the compensation packages of DoD employees compare favorably with those of contractor employees. However, flexibilities in contractor personnel systems open up possibilities for much higher earnings if an employee is willing to work hard, take risks, and deliver results. From a more sociological perspective, retention of employees is aided by a positive, mission-focused culture, which can be created and sustained both in contractor and DoD organizations provided that strong, effective leaders are present. In conclusion, to motivate strong performance, organizations have a variety of tools from which to choose. These tools transcend base salary to encompass a shared sense of purpose, positive morale, leadership development, mentoring, and fresh ways of thinking about career progressions. When reinstituting the GS system, the DoD should fully leverage the management flexibilities and workforce incentives (toward these ends), as authorized by Congress.

References

- Allison, G. (1979, November 19-20). Public and private management: Are they fundamentally alike in all unimportant respects? In *Proceedings of the Public Management Research Conference* (pp. 27-38); (1980, February). OPM Document 127-53-1. Washington, DC.
- Allison, G., & Szanton, P. (1976). *Remaking foreign policy: The organizational connection*. New York: Basic Books.
- Asch, B. (2005). The economic complexities of incentive reforms. In Klitgaard & Light (Eds.), *High-performance government: Structures, leadership, incentives* (pp. 309-342). Santa Monica: RAND.
- Baker, G., Jensen, M., & Murphy, K. (1988). Compensation and incentives: Practice vs. theory. *Journal of Finance*, 43, 593–616.
- Blair, M.M., & Kruse, D.L. (1999, fall). Worker capitalists? Giving employees an ownership stake. *Brookings Review*, 17.



- Borjas, G. (2002, October). *The wage structure and the sorting of workers into the public sector* (Working Paper 9313). Cambridge, MA: National Bureau of Economic Research.
- Bozeman, B. (1987). *All organizations are public*. Jossey-Bass.
- Branham, L. (2001). *Keeping the people who keep you in business*. New York: American Management Association.
- Branham, L. (2005). *The seven hidden reasons employees leave*. New York: American Management Association.
- Defense Finance and Accounting Service. (2009). *Military pay tables*. Retrieved from <http://www.dfas.mil/militarypay/militarypaytables/2009MilitaryPayTables.pdf>
- Eisenhardt, K.M. (1989, January). Agency theory: An assessment and review. *Academy of Management Review*, 14(1), 57-74.
- GAO. (2009). *Additional data and actions are needed to effectively manage and oversee DoD's acquisition workforce*. Retrieved from <http://www.gao.gov/new.items/d09342.pdf>
- Gellerman, S.W. (1992). *Motivation in the real world*. New York: Penguin Group.
- Gibbs, M. (2001). *Pay competitiveness and quality of Department of Defense scientists and engineers* (MR-1312-OSD). Santa Monica, CA: RAND.
- Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3), 481-510.
- Hedgpeth, D. (2009). Pentagon to expand its acquisition force. *Washington Post*. Retrieved from http://voices.washingtonpost.com/government-inc/2009/04/pentagon_to_expand_its_acquisi.html
- Hirsch, P., Michaels, S., & Friedman, R. (1987). 'Dirty hands' v. clean models: Is sociology in danger of being seduced by economics? *Theory and Society*, 317-336.
- Holmstrom, B., & Milgrom, P. (1991, January). Multitask principal-agent analyses: Incentive contracts, asset ownership and job design. *Journal of Law, Economics and Organization*, 7(Special Issue: Papers from the Conference on the New Science of Organization), 24-52.
- James, K.C. (2002). *A fresh start for federal pay: The case for modernization*. White Paper. Washington, DC: US Office of Personnel Management.
- Jamieson, D., & O'Mara, J. (1991). *Managing workforce 2000: Gaining the diversity advantage*. San Francisco: Jossey-Bass.
- Katz, L., & Krueger, A. (1991). *Changes in the structure of wages in the public and private sectors* (Working Paper No. 2667). New York: National Bureau of Economic Research.
- Lazear, E.P. (2000). Performance pay and productivity. *American Economic Review*, 90(5), 1346-1361.
- Lee, S., & Whitford, A. (2008). Exit, voice, loyalty and pay: Evidence from the public workforce. *Journal of Public Administration Research and Theory*, 18(4), 647-671.



- Miller, G., & Whitford, A. (2007). The principal's moral hazard: Constraints on the use of incentives in hierarchy. *Journal of Public Administration Research and Theory*, 17(2), 213-233.
- Roberts, N.C. (1991). *Limitations of strategic management in bureaus: The case of the Department of Defense* (Working Paper). Naval Postgraduate School. Retrieved from <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA236803&Location=U2&doc=GetTRDoc.pdf>
- US Congress. (2004). *National defense authorization act* (P.L. 108–136). Retrieved from <http://www.dod.mil/dodgc/olc/docs/2004NDAA.pdf>
- US Congress. (2010). *National defense authorization act for 2010* (P.L. 111-84). Retrieved from http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=111_cong_bills&docid=f:h2647enr.txt.pdf
- US Office of Personnel Management. (1998, August). *HRM Policies and Practices in Title-5 Exempt Organizations* (MSE-98-4). Washington, DC: Office of Merit Systems Oversight and Effectiveness.
- US Office of Personnel Management. (2001, July). Human resources flexibilities and authorities in the federal government. Washington, DC: Office of Merit Systems Effectiveness, Center for HR Innovation.
- US Office of Policy and Evaluation, Merit Systems Protection Board. (2001, December). *The federal merit promotion program: Process vs. outcome*. Washington, DC: US GPO.
- Wilson, T.B. (1994). *Innovative reward systems for the changing workplace*. New York: McGraw-Hill.



Industry Perceptions of Department of Defense Program Manager Competencies

Roy Wood—Dr. Wood is currently the Dean of the Defense Systems Management College, School of Program Managers at the Defense Acquisition University at Fort Belvoir, VA. He previously served as the Principal Assistant Deputy Under Secretary of Defense for International Technology Security under the Director, Defense Research and Engineering. He is a retired Naval Engineering Duty Officer with 24 years in operational and acquisition tours and has worked in the defense industry.

Roy L. Wood, Jr., PhD
Dean, Defense Systems Management College, School of Program Managers
Defense Acquisition University
9820 Belvoir Road
Ft. Belvoir, VA 22060
(703) 805-4558 (w)
(540) 371-1554 (h)
roy.wood@dau.mil

Abstract

Large, complex Defense Department weapons system acquisition programs have been plagued by cost overruns, delayed schedules, and subpar performance. Much of the blame has been placed on government program managers (PMs). This study provides a new perspective on government PM competencies by surveying defense industry managers who work with the government PMs on a day-to-day basis. 146 industry managers rated the importance of PM competencies and assessed how well, from their perspective, their government PM counterparts met those competencies. The data gathered from this survey revealed several surprising insights, including a conclusion that government program managers' performance on several key *technical* skills may need improvement. The results of this study may be useful in updating training and development strategies for government PMs to improve program outcomes.

Introduction

The US Department of Defense uses a program management structure to acquire its sophisticated sea-, land-, air-, and space-based systems. Under this management paradigm, a civilian or military program manager, leading a team of government engineers, logisticians, business and financial managers, contracting officers, and administrative personnel, is responsible for the development and delivery of his or her system. The GAO reported in 2005 (p. 14) that there were 729 program managers executing programs in the DoD.

In a typical defense project, the government team works closely with representatives of the operational (warfighter) community to understand new system needs and requirements and to translate these needs into performance or technical specifications. The government team is responsible for cost and schedule estimates for the program, as well as for describing the procurement and contracting approaches, test and evaluation plan, and the strategy for supporting the system over its lifecycle. These plans and strategies evolve within a complex bureaucracy of checks and balances that provides oversight and assistance in creating the program.



After plans are in place for the acquisition, the government program team creates a contract solicitation and accepts proposals and bids from companies in the defense industry. After a contract is awarded, the government team works closely with its industry partner to proceed to system development. This process is typically slow and methodical, proceeding through a series of decision milestones in which progress is gauged by an oversight official above the program manager's organizational level. Throughout the development program, the government PM must advocate and negotiate for his or her program's funding, requirements, program scope, schedule, personnel, and myriad other details. The PM must also work closely with the contractor to evaluate alternative technologies and industrial processes, monitor contractor spending and adherence to program schedule, evaluate progress and quality of workmanship, and provide general oversight to protect the government's equities. Since the end of the Cold War, defense acquisition has undergone significant transformation and downsizing, and the government relies more heavily on contractors for detailed designs to meet program objectives (Nissen, Snider & Lewis, 2002). This has created greater partnering and closer relationships between government program teams and their industry counterparts (Jones, 1997).

The government program manager must be technically competent, able to manage technology and system engineering as well as software and information systems, and understand manufacturing and industrial processes. He/she must demonstrate key business competencies such as financial management, contracting, and cost estimating. The PM must exercise management acumen in developing and executing the program strategy, managing core processes, and dealing with the day-to-day management challenges of a large, complex program. The PM must also exhibit leadership competencies when leading blended government-industry teams or engaging in negotiations and advocacy with customers and stakeholders. Fox and Miller (2006) summed up the management challenge this way:

Managing [a large complex project] is more than a science; it is a continually evolving art... Managers must augment a strong foundation of conventional management skills in planning, organizing, and controlling, with knowledge of the requirements, resources, and constraints of a specific project as it progresses. (p. 109)

Purpose of the Study

Given that PMs must possess an expansive portfolio of required competencies, it is natural to ask whether some competencies may be more important in helping to assure program success. For example, research by Bauer (2006) asserted that *management competence* is perceived to be more important for defense and aerospace industry program managers than technical skills. One might assume that the same would be true for government PMs in the same industry, and this is one key aspect that this study sought to examine.

The purpose of this study, then, was to (a) determine the relative importance of key competencies of government PMs and (b) assess PM performance against these competencies. In this way, it would be possible to rank-order competencies according to their perceived importance to program success and to compare these to PM performance. In this way, a reasonable path forward might be identified to help strengthen PM competencies through training and development.

Most PM competency studies in the literature, however, are either self-studies in which PMs rank-order the competencies and rate their own performance against them, or



they are studies in which immediate supervisors of the program managers are surveyed. This study took a different approach. Leveraging the close working relationship between the government and industry program manager, this study captured opinion data from industry managers to provide a more objective *peer assessment* and to avoid potential “blind spots” in the self-assessed competency data contained in the current literature.

Research Questions

The research questions for this study were:

1. Which project management competencies are perceived by industry managers as most important in government program managers?
2. How well are government program managers perceived by their industry partners to be meeting those competencies?

Competency Theory and Program Management Competencies

The roots of competency theory originate in Frederick Taylor's studies of scientific management, breaking complex tasks into manageable constituents, and assigning specialized workers using standardized manufacturing processes to each task to increase efficiency and effectiveness (1911). Implicit in Taylor's study was the assumption that the more competent the worker becomes at each specific work task, the more successful the overall job will be. These work studies, however, tended to be highly detailed and task-oriented, rather than focused on the skills and abilities of the individual worker. Indeed, optimizing tasks and training individuals to become specialists, as with a manufacturing assembly line, may create a workforce with little flexibility or incentive for creativity (Womack, Jones & Roos, 1991).

The search for competence is more closely focused on the individual and attempts to identify the ideal attributes of a high-performing worker (McClelland, 1973). Early efforts to understand what makes individuals competent focused primarily on human intelligence factors. It was reasoned that the more intelligent a person is, the better he or she will learn, adapt, and perform on any given job. Research on intelligence testing and improvement ensued, but in the end, yielded few real gains for organizations (Berger & Berger, 2003).

Harvard psychologist David McClelland was one of the first to argue effectively against the prevailing intelligence-centric view of job competence. He is considered by many to be the pioneer of modern competency theory (Draganidis & Mentzas, 2006; Garman & Johnson, 2006; Ruth, 2006). In his seminal work, McClelland (1973) argued that competence—work-related knowledge, skills, behaviors, and abilities—was a better practical predictor of superior work performance than was intelligence. McClelland further argued that intelligence tests were discriminatory and therefore largely invalid, “favoring certain ethnic and socioeconomic groups” (Gale, 2007, p. 143). McClelland suggested the more direct approach of testing job performance against required skills, which is the approach still used in most modern competency methodologies.

However, universal definitions of *competence*, *competency*, and *competencies* are difficult to find in the literature and are not generally agreed upon (Whiddett & Hollyforde, 2003). Gale (2007) argued that defining competence is difficult because “competence is a normative concept rather than a descriptive one” (p. 145). He went on to say that competence is being “concerned with the capacity to undertake specific types of actions, and it can be considered a holistic concept involving the integration of attitudes, skills,



knowledge, performance, and quality of application” (p. 145). In a similar, but more practitioner-focused vein, Parry (1998) defined competency as “a cluster of related knowledge, attitudes, and skills that affects a major part of one’s job (i.e., one or more key roles or responsibilities); that correlates with performance on the job; that can be measured against well-accepted standards; and that can be improved via training and development” (p. 60).

Operationally, managing competence in an organization involves identifying a set of highly desirable attributes that can positively influence desired organizational outcomes. Researchers and practitioners seek to create competencies and competency models that can be used to influence hiring, retention, and training practices to improve the quality of the organizational workforce. To meet this desire, the most common working definitions in the literature have included some aspects of technical, social or interpersonal, and cognitive or problem-solving that, ideally, can be measured and improved over time. In practice, many organizations, like the American Society of Training and Development (Bernthal et al., 2004, p. xix), and the Project Management Institute (PMI, 2002) have adopted working definitions relating individual behaviors with job performance, which they believe can be of value as the starting point for identifying specific competency-related attributes and developing effective competency models.

Program Manager Competencies

Project management is a relatively new field of professional endeavor. Much of the research into those things that make projects successful is also relatively recent. Crawford (2006) summarized this brief history:

The first signs of project management as a distinct field of practice were the network analysis and planning techniques, like PERT and CPM, that emerged in the 1950s for use on major projects in construction, engineering, defense, and aerospace industries. Users of these tools and techniques recognized shared interests leading to the formation of project management professional associations in the 1960s, initially to facilitate knowledge sharing between practitioners. The mid-1990s were a crucial point in the development of project management standards and related certification programs (p. 75).

In the United States and many international locations, project manager standards and certifications have been developed by the Program Management Institute (PMI). The PMI identified the knowledge, skills, abilities, and behaviors needed to be an effective project manager through its Project Management Competency Development Framework (PMI, 2002). This framework operationalizes Parry’s definition of competencies and describes PM competencies along three dimensions: knowledge (what PMs know), performance (what PMs do with the knowledge), and personal competency (how PMs behave—their attitudes and personality traits) (PMI, 2002). The overall approach and goal of the PMI is consistent with Toney’s (2002) view that “application of validated competencies assures project customers and stakeholders that the probability of project success is improved” (p. xix). Crawford (1997) advanced the idea that there is a causal relationship between PMI competencies and job performance, lending a strategic notion to competency identification and development.



The Need for Competent Defense Program Managers

For the past several decades, news reports of \$600 toilet seats, poor performance of battlefield equipment, and cancelled programs have been all too commonplace (Besselman, Arora, & Larkey, 2000; Samuel, 2003). The Defense Acquisition Program Assessment (DAPA) Report of 2006 asserted that:

Both Congress and the Department of Defense senior leadership have lost confidence in the capability of the Acquisition System to determine what needs to be procured or to predict with any degree of accuracy what things will cost, when they will be delivered, or how they will perform. (Kadish, p. 1)

“Program manager’s expertise” was identified in the report as one of the top five issues contributing to the poor performance (Kadish, p. 3). In 2008, there was sharp criticism of both poor performance of the defense acquisition system and the program managers who run those acquisition programs. The Government Accountability Office (GAO), an independent research arm of the US Congress, commented recently that the DoD needed to “strengthen training and career paths as needed to ensure program managers have the right qualifications for running the programs they are assigned to” (Sullivan, 2008, p. 16).

Improving program manager competencies is believed to be essential to improving the overall performance of the defense acquisition system. Identifying key competencies that drive program performance and specific shortfalls in program manager knowledge, skills, and abilities is an important first step and the aim of this study.

Research Methodology

The independent variables in this study were the competencies of government PMs. These included 20 technical (hard) skills and 15 behavioral (soft) skills, as described below:

Technical Project Management (Hard Skills) Competencies (C1-20)

- C1. Determine program goals. The ability to work with program stakeholders in order to understand the program's requirements and specifications.
- C2. Determine program deliverables. The ability to work with program stakeholders to generate a scope of work, requirements, and/or specifications for the program.
- C3. Technical ability. The ability to understand and be conversant in the core technologies of product/deliverables of the program.
- C4. Document program constraints. The ability to lead the program team to uncover and document possible program constraints that could affect program completion.
- C5. Document program assumptions. The ability to lead the program team to determine information that must be validated or situations that must be controlled during the program process in order to facilitate program planning.
- C6. Define program strategy. The ability to evaluate possible strategies or alternative approaches to meet the program's requirements and/or specifications.



- C7. Quality assurance. The ability to identify performance criteria using product/service specifications, technical expertise, and standards to ensure performance standards are met, customer expectations are met, and processes are analyzed for further improvements.
- C8. Identify resources requirements. The ability to identify key resource requirements needed to support planning and decision-making.
- C9. Develop a budget. The ability to complete cost estimates and produce a program budget to support planning and decision-making.
- C10. Create a work breakdown structure (WBS). The ability to use the scope of work and other project documents to develop a work breakdown structure to facilitate project planning.
- C11. Develop a schedule. The ability to complete a program schedule that supports planning and decision-making.
- C12. Develop a resource management plan. The ability to develop and publish a resource management plan (human resources, procurement, etc.) by identifying resource requirements and obtaining commitment from internal and external assets that enable completion of all program activities.
- C13. Establish program controls. The ability to establish program controls by establishing targets and plans, measuring actual performance, comparing actual performance against planned performance, and taking necessary actions to correct the situation.
- C14. Develop program plan. The ability to develop a formal comprehensive program plan documenting deliverables, acceptance criteria, process, procedure, and tasks to facilitate program completion.
- C15. Communicate program status. The ability to produce program reports and presentations that provide timely and accurate program status and decision-support information to upper management, customers, and fellow team members.
- C16. Measure program performance. The ability to compare actual results to a documented baseline in order to identify program trends and variances.
- C17. Implement corrective action. The ability to take timely corrective action by addressing the root causes in the problem areas in order to eliminate or minimize negative impact to the program.
- C18. Implement change control. The ability to track and document all potential improvements and other changes in scope, specifications, cost, or schedule and analyze the consequences of these changes in relation to the overall project.
- C19. Respond to risk. The ability to respond quickly to risk event triggers in accordance with the risk management plan in order to keep the program on schedule and within budget.
- C20. Conduct administrative closure. The ability to conduct financial closure and publish formal program closure documents.



Personal (Soft Skill) Competencies (CS1-15)

- CS1. Project leadership. The ability to set a vision, identify the action steps, and motivate others to maintain their commitment to program success. The ability to influence a team to willingly work toward predetermined program objectives.
- CS2. Flexibility. The ability to adapt and deal with situations and manage expectations during periods of change and uncertainty during a program.
- CS3. Sound business judgment. The ability to stay focused on the business target. The program manager knows the organization's business purpose of program and makes decisions within that context.
- CS4. Trustworthiness. The ability to build positive working relationships and credibility with team members, upper management and stakeholders.
- CS5. Communication style. The ability to adapt one's communication style to fit the situation and the audience. The ability to present information without bias and exchange information in a clear and unambiguous manner.
- CS6. Listening Skills. The ability to ensure all team members have a chance to provide input to the program. The ability to read body language and perceive group dynamics.
- CS7. Setting and managing expectations. The ability to communicate with all program stakeholders, especially customers, and address program objectives, timelines, budgets, risks, and estimates. The ability to clearly communicate program changes and/or adjustments with support rationale to the customer in a proactive manner.
- CS8. Negotiations. The ability to develop win-win situations that culminate with both parties being satisfied with the final agreement.
- CS9. Issue and conflict resolution. The ability to understand and implement conflict resolution models for resolving issues and preventing the conflict from affecting the program's outcome.
- CS10. Organizational skills. The ability to arrange program activities in such a way that they systematically contribute to the program's goals.
- CS11. Coaching. The ability to provide feedback to team members and stakeholders in a positive manner that builds trust and credibility.
- CS12. Facilitation. The ability to facilitate or guide team members through a process that helps them discover answers and overcome barriers to successful program completion.
- CS13. Decision making. The ability to make the best choice from among many alternatives.
- CS14. Problem solving. The ability to identify issues, to conduct accurate assessments of the issues, and propose viable solutions to issues.
- CS15. Team building. The ability to encourage and enable people to work together toward a common goal



Competency Survey

An online survey questionnaire was developed for this study that consisted of three parts: (a) ratings for technical (hard) skills, (b) ratings for management/leadership (soft) skills, and (c) demographic questions about the participant and program. The survey used in this study was a modification of one used by Golob (2002) to identify competencies that might be useful in hiring or promotion decisions for project managers. The modified survey instrument was subjected to an expert review. Pilot study data were validated through item analysis, and Cronbach's alpha tests.

The survey instrument asked industry participants to assess the importance of the 20 technical and 15 management (soft-skill) competencies to program success, and then to assess how well, in his or her judgment, their government counterpart performed in those competencies. Each of the responses was based on a 5-point Likert-type scale. Participants were asked to assess the importance of each competency to program success with ratings from 1 (indicating that competency is *unimportant/not needed*) to a value of 5 (indicating that the competency is *extremely important*). In assessing performance against each competency, the participant rated each from 1 (indicating that the government PM is *not meeting the competency*) to a value of 5 (indicating that the PM is *working at an expert level*). Four additional demographic questions were asked about the study participant and his/her program experience.

Survey Responses

Using the 2005 GAO (p. 14) estimate of 729 programs in the Department of Defense, an appropriate sample size was calculated. Alreck and Settle (1995) suggest a non-probability sample of about 10% of the parent population, or approximately 73 respondents, based on the estimated population. Another, more conservative, approach was to calculate the number needed for a probability sample. In this case, assuming a 95% confidence level with a $\pm 10\%$ confidence interval, as shown in the calculation below, the minimum sample size would be 85 participants.

To protect the anonymity of participants and their companies, senior executives from several defense industry corporations were asked for assistance in referring potential study participants in their companies who were managing defense programs. Since, by some estimates, only about a 30% response rate could be expected from Internet surveys (Sue & Ritter, 2007), each executive was asked to refer 75 to 100 program managers to the online survey to assure a sample size of at least 85 participants. In all, 146 surveys were completed, well exceeding the original 85 target. An additional 71 were started but abandoned, and 83 were started but not completed by the survey closeout date.

Demographics

The survey asked four questions to help understand the study participants and the programs they managed:

1. How many years experience do you have as a program/project manager?
2. What is the Acquisition Category (ACAT) rating of your program?
3. What is the acquisition phase of your program?
4. On average, how often do you communicate with the government Program Manager (face-to-face, telephone, e-mail, other)?



Over 78% of participants reported that they had 10 or more years of experience, and nearly half (48%) reported that they managed some of the largest, most complex programs (ACAT I or II) in the Department of Defense. These responses indicate that the participants were highly experienced project managers with significant responsibility for managing challenging programs, thus they should be expected to have a good practical understanding of the competencies involved in complex program management.

Almost half the participants indicated that their programs were beyond development and into the later, more mature production and deployment phases. When asked to rate the frequency of interaction with their government counterparts, over three quarters of participants indicated that they communicated with their government counterparts often, very often, or daily, indicating that responses were generally well informed. Demographic responses are shown in Table 1 below.

Table 1. Survey Demographics

Category	Variable	N	%
Experience Level in Years	15 or more	64	44.1
	10-14	50	34.5
	5-9	19	13.1
	0-4	12	8.3
Program Acquisition Category ACAT	ACAT I	54	37.2
	ACAT II	16	11.0
	ACAT III	7	4.8
	Other	68	46.9
Program Phase	Production & Deployment	71	49.0
	Sys Design & Development	53	36.6
	Technology Development	19	13.1
	Concept Refinement	2	1.4
Communication Frequency with Government PM	Daily	43	30.1
	Very Often	48	33.6
	Often	21	14.7
	Occasional	19	13.3
	Infrequent	12	8.4

Research Question 1

Survey data were analyzed to address the first research question: “Which project management competencies are perceived by industry managers as most important in government program managers?” Participants responded to a list of 20 technical competencies (C1 through C20) and 15 soft skill competencies (CS1 through CS15), rating the relative contribution of each to program success. Each competency was listed on the questionnaire with Likert-type scale choices of *Very Important*, *Important*, *Neutral*,

Unimportant, or Very Unimportant. Figure 1 shows the average scores for each competency.

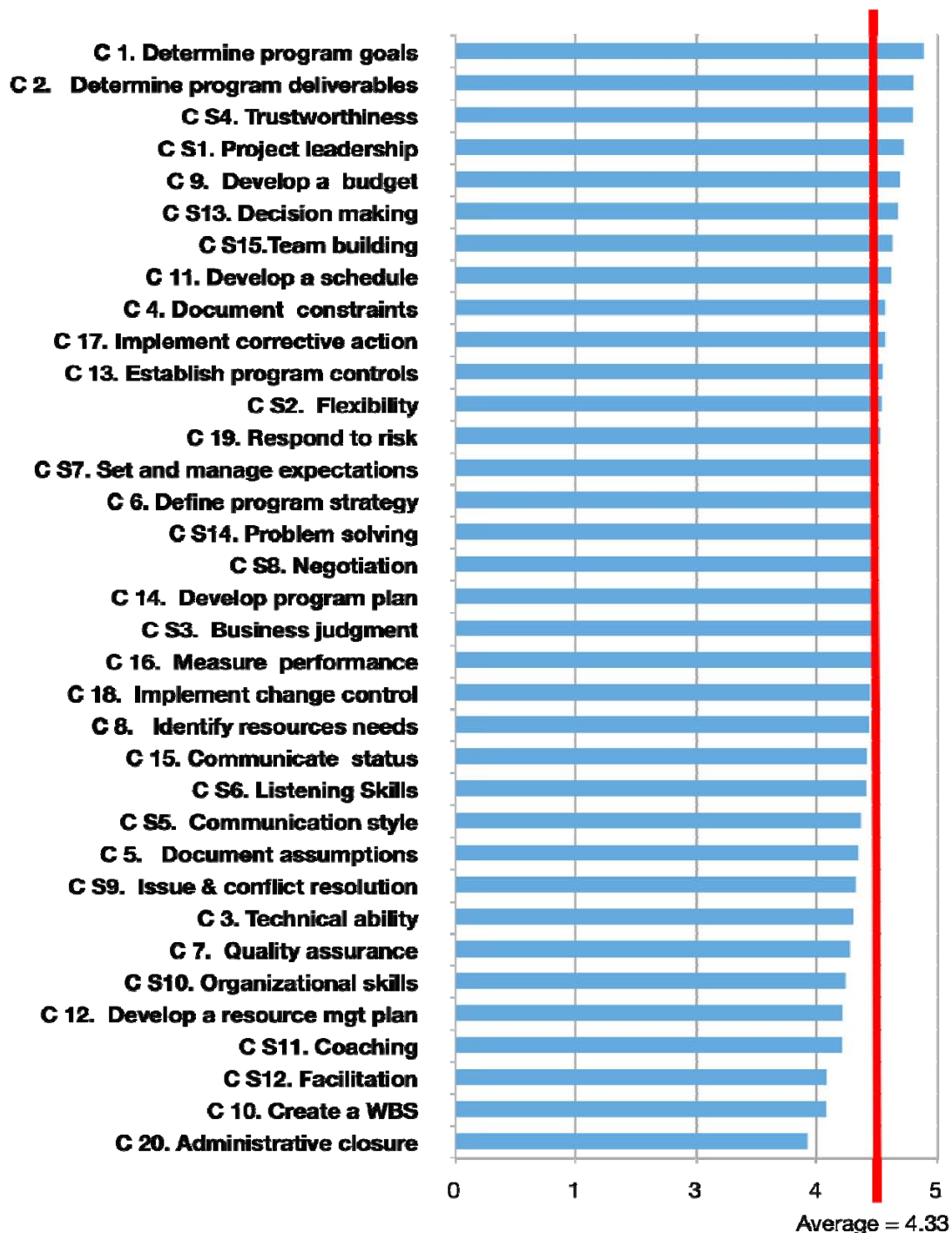


Figure 1. Ranking of Competency Importance

The data show that most of the competencies scored quite high for their contributions to program success (mean = 4.33). The high scores affirmed the selection of appropriate competencies that most influence program success, but the close rankings limit the ability to clearly identify only a few competencies from the list that have the greatest impact.

The highest-rated competencies represented a mix of technical and soft skills. The most valued technical skills were the ability to determine program goals and deliverables and to develop a program budget. Such results are not surprising. Among others, Pinkerton (2003, p. 53) pointed out that the first criterion for project success is to have clearly defined goals and objectives. It is important for the government to specify the deliverables from the project, and it is equally important for industry, because deliverables define the government's expectations in concrete terms. Similarly, a sound program budget is important to match resources to goals and deliverables. Fox and Miller (2006) observed how these skills must be related:

Managing a [large, complex project] is more than a science; it is a continually evolving art. ...managers must augment a strong foundation of conventional management skills in planning, organizing, and controlling, with knowledge of the requirements, resources, and constraints of a specific project as it progresses. (p. 109)

The most highly rated soft skills included trustworthiness, project leadership, and decision-making. The importance of trust to proper organizational and inter-organizational functioning has been widely documented in the literature (Jehn & Mannix, 2001; Joseph & Winston, 2005; Wells & Kipnis, 2001). Trust may be particularly important in large, complex projects in which not every expectation can be instantiated in the government-industry contract. Trust and understanding between the government and industry managers are essential to minimizing conflict, fostering cooperation, and succeeding.

In a complex defense project, the government program manager must be the leader who sets the vision and goals, motivates the team, and is committed to program success. This role cannot be assumed by the industry manager or by any other member or stakeholder in the program. Related to this idea, the government PM must be willing and capable of making and influencing the myriad daily decisions and choices that shape the outcomes of a program. Imparting a good deal of wisdom, Fox and Miller (2006) noted, "Skilled project managers focus more on monitoring and influencing decisions, and less on giving orders. Clearly, project managers have substantially more responsibility than authority" (p. 124).

In such a complex environment that requires delicate balancing of program goals and powerful stakeholder interests, it is unlikely—perhaps impossible—for program manager decisions to be entirely based on rational, stepwise decision making aimed at clear-cut outcomes. Rather, it is more likely that program manager decision-making is better explained by behavior theory (Cyert & March, 1992). In this theory, complex decisions are the outcome of organizational behavioral factors such as quasi-resolution of conflicting program goals and avoidance, when possible, of uncertainties that create program risk.

Program Complexity and Experience

One question arising from the data was whether or not the program complexity or the experience level of survey participants might affect the ranking of competencies for importance. Programs that are higher in cost and complexity are generally considered to be



more challenging to manage. Program management teams for more complex programs are typically larger, with more stakeholders who have greater influence. Larger programs are subjected to higher levels of oversight and scrutiny; they are required to do more formal planning and documentation and are subject to more frequent and detailed reports on status. Program managers of large, complex programs have a greater scope of responsibility and a substantially larger span of control, and, conceivably, these program managers require a different set of skills and competencies.

To determine whether program cost and complexity influenced the ranking of competencies for importance, an additional test was performed on the data. A *t*-test was performed on the competency importance means to examine the responses of managers of highly complex ACAT I programs in comparison to those of participants in lower acquisition category programs. The results showed no significant statistical differences (95% confidence level) between the perceptions of these two groupings of industry program managers.

Similarly, the experience level of program manager participants could have influenced the judgment used in rating the importance of competencies. A *t*-test was conducted on the dataset by splitting it between participants with more than 10 years of experience and those with less. As with the ACAT level, experience level seemed to have little effect on perceived competencies, with two exceptions.

From the *t*-test, experience seemed to influence the rating of competencies CS5 (Communication style) and CS13 (Decision-making). Communication style was more highly rated by less experienced participants with a 4.39 mean, versus a 4.16 mean for participants with greater than 10 years of experience. Conversely, decision-making was rated more highly by more experienced participants with a mean score of 4.65. Less experienced participants rated this competency 4.39.

The importance of decision-making skills in a government program manager is critical to program success. That this skill is more highly valued by experienced industry managers is an insightful finding. Similarly, the ability to exchange information in a meaningful way, especially in the collaborative environment of a government-industry project, seems to resonate with less experienced managers who may feel a greater need or appreciation for the government PM to clearly articulate program direction. It may be noteworthy that improving communication between government and industry was recognized as a key factor in program success and one of the motivating forces behind the Department of Defense acquisition reform initiative of the late 1990s (Jones, 1997).

Research Question 2

Survey data were also collected to answer the second research question: “How well are government program managers perceived by their industry partners to be meeting those competencies?” The questionnaire asked participants to respond to each of the competencies with their assessment of how well their government PM counterpart met the competency. The Likert scale observations included ratings of *Expert*, *Good*, *Average*, *Fair*, *Poor*, and a no-response choice. Figure 2 summarizes the rank-ordered response for each competency.



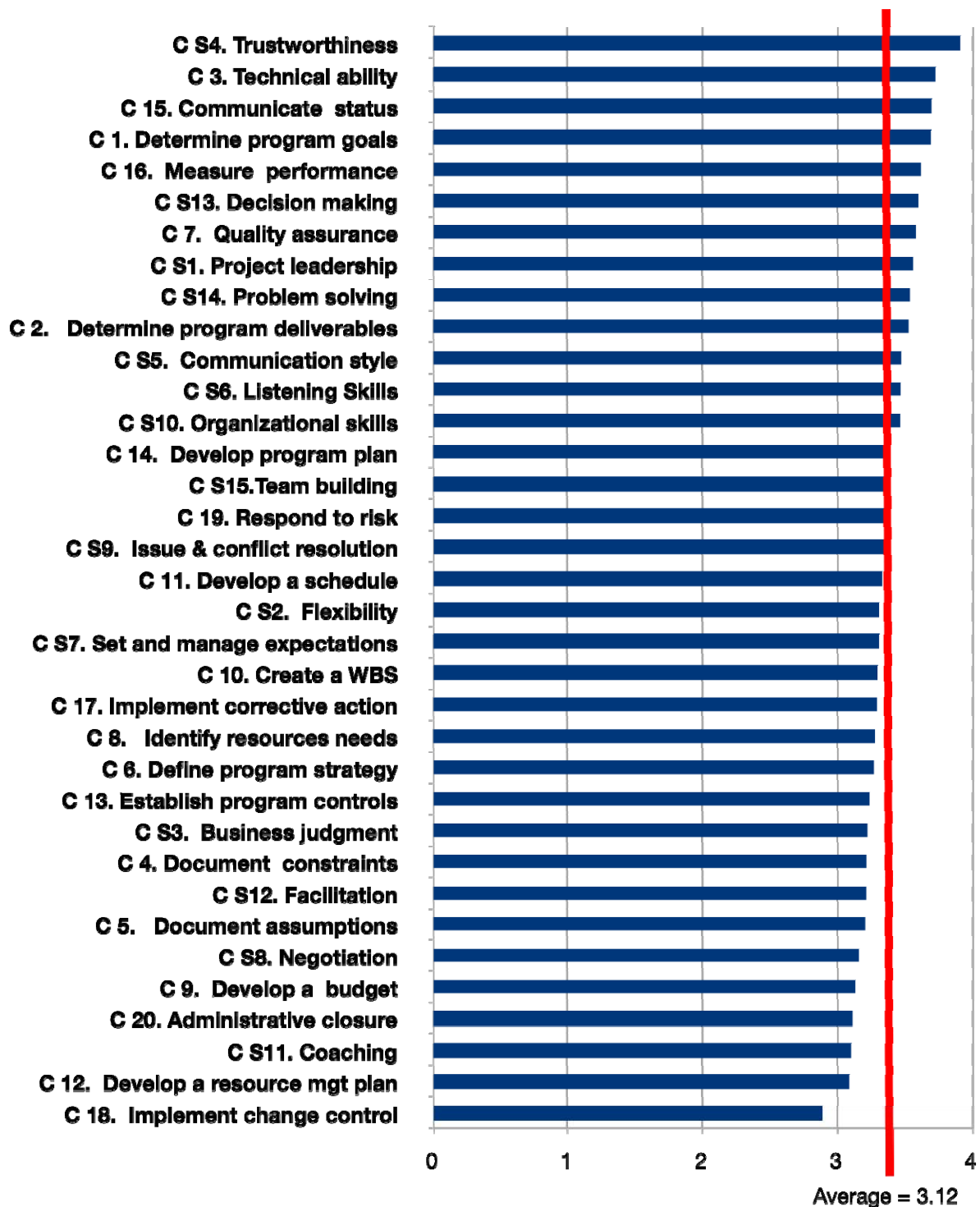


Figure 2. Ranking of Competency Performance

It is noteworthy that the average from all the participants rates government PM performance between *fair* and *good*, with most nearer *average* performance. This was disappointing, given the high stakes and expectations for managing billions of taxpayer dollars to provide critical defense systems to the battlefield.

Sensitivity to Program Phase and Contact Frequency with Government PM

To help understand whether the performance assessment responses provided by the industry participants may have been influenced by the program phase or the frequency of contact with government counterparts, additional statistical tests were conducted. It was considered possible that the phase of the program could have limited whether participants had the opportunity to observe the government PM performing some activities.

For example, programs in production and deployment may have fewer planning activities required, such as developing a work breakdown structure or schedule, and less opportunity for these skills to be observed and assessed. Similarly, and perhaps more intuitively, the frequency of interaction between the government program manager and the industry counterpart could have affected the assessment of competencies. An industry manager who rarely worked directly with a counterpart may not have had the exposure to be able to formulate accurate assessments. Conversely, a situation providing closer observation could result in scores that are more reliable.

To assess these eventualities, *t-test* calculations were used to check the competency performance against the participants' program phase and frequency of communication. A *t-test* of competency performance means versus program phase was performed first. The data set was bifurcated into (a) production and deployment (PD) and (b) data from earlier design and development intensive phases. In these two groupings, 67 participants reported programs in production and deployment, while 71 reported programs in earlier phases. The results showed no significant statistical differences (95% confidence level) between the perceptions of these two groupings of industry program managers in these different program phases.

Similarly, a *t-test* was performed to determine if the performance means differed significantly with frequency of communication between industry participant and government PM. The data set was bifurcated between those who responded *Daily* or *Very Often* and those who had less frequent contact with their government counterpart. Here, too, there were no significant differences among responses from those who had daily or frequent contact with their government counterparts and those who had less contact, with two exceptions.

For competency C14 (the ability to develop a program plan), those participants with more frequent contact had a higher average score of 3.30, versus 2.83 for those participants who communicated with their government counterpart less frequently. The other competency, CS15 (team building), differed, with participants who said they had more frequent contact with their government PM counterparts ranking this competency higher, with an average of 3.26 versus 2.89 for those participants reporting less frequent contact. Possible explanations of these differences are discussed in Chapter 5.



Competency Performance Shortfall

To answer the second research question in a manner that could provide insights for improving PM competency, the *performance shortfall* was determined. The most simplistic method used was to examine the lowest-rated competencies without regard to perceived importance. To do this, the difference was calculated between the best possible rating of 5.0 and the average reported rating for each competency from the survey. By this measure, the top 10 competencies needing improvement are shown in Table 2 below.

Table 2. Competency Performance Shortfall

Competency		Rating	Shortfall 5.0-Rating
C 18	Implement change control	2.676	2.324
C 12	Develop a resource mgt plan	2.860	2.140
C S11	Coaching	2.874	2.126
C 20	Administrative closure	2.884	2.116
C 9	Develop a budget	2.902	2.098
C S8	Negotiation	2.927	2.073
C 5	Document assumptions	2.971	2.029
C S12	Facilitation	2.977	2.023
C 4	Document program constraints	2.988	2.022
C S3	Business judgment	2.985	2.015

The downside to simply choosing the lowest-rated competencies, however, is that such a method does not consider the perceived importance of each competency toward program success. The competency *implement change control*, for example, ranked lowest in performance but was considered only moderately important. Similarly, while *administrative closure* ranked fourth lowest in performance, it was also last in importance. In identifying competencies for possible improvement, not only should performance itself be considered, but performance in relation to importance should be heeded as well. This was the aim of the next calculation.

To ensure that the competency *importance* was considered more heavily when ranking competency shortfalls, a more complex computation was used to add more weight to competency importance. Employing a weighting model used by Borich (1980), the difference between competency importance and competency performance was multiplied by the mean competency importance. In other words:

Step 1: mean competency importance - mean competency performance = discrepancy

Step 2: discrepancy × mean competency importance = weighted discrepancy score

The top-10 list derived from the Borich model is shown in Table 3. Using this method of weighting the shortfalls creates a different ranking of the competencies, favoring those with higher importance scores. Note that in this ranking, *develop a budget* ranked first because of its high importance ranking and low performance score.



Table 3. Competency Shortfalls Using Borich Model

	Competency	Importance (I)	Performance (P)	Difference I x (I-P)
9	C Develop a budget	4.616	2.902	7.913
2	C Determine program deliverables	4.753	3.268	7.060
18	C Implement change control	4.308	2.676	7.030
1	C Determine program goals	4.863	3.420	7.016
4	C Document program constraints	4.466	2.978	6.643
11	C Develop a schedule	4.527	3.088	6.519
13	C Establish program controls	4.438	3.000	6.384
S15	C Team building	4.538	3.132	6.378
S8	C Negotiations	4.377	2.927	6.345
17	C Implement corrective action	4.466	3.051	6.316

In this list, a surprising number of technical skills topped the list, including *develop a budget*, *implement change control*, *document program constraints*, and *determine program deliverables*. Of the top 10 items, only two identified shortfalls were soft skills, and even these were near the bottom of the list: *negotiation* and *team building*.

Conclusions

The current study appeared to be the first in the literature to explore the competencies of Department of Defense program managers from the perspective of their industry counterparts. The data allowed for the ranking of competencies believed to contribute most importantly to program success and for an assessment of how well defense program managers met those competencies. From these results, a priority-ordered list was developed of competencies that could be candidates for improvement through training and development. The competencies ranking in the top 10 for importance represented a relatively even mix of technical and soft skills, as did the raw rankings of PM performance.

However, when analysis was done to discover the variance between competency importance and performance, the results ranked many of the technical skills at the top of the list of candidates for improvement. This appeared contrary to assertions by Bauer (2006) and Golob (2002) that soft skills may be the most important to program success, and it contradicted Gadeken (2004), who suggested that defense PMs should seek soft-skill training. These findings seem to refute the conventional wisdom and may provide a valuable new insight and contribution to the literature.



Practitioner Application

The implications of this study are clear for the practitioner. From the data, there appears to be a need for greater technical training and development in government program managers. The specific competency shortfalls identified in the analysis may be helpful in directing program managers and the training establishment toward specific skills that would benefit most from remediation.

By extension, since government program managers depend upon their functional staffs for support in many of the technical skill areas, the current study may point toward shortfalls within those functional areas. For example, to help remedy the highest ranked shortfall, budget development, additional training may be needed in the business and financial management career field to provide better support to the program manager in this critical area. This may include instruction and practice on developing detailed program cost estimates, allocating budget appropriately to critical tasks in the work breakdown structure, and finding ways to respond to funding needs for emergent or evolving program changes.

References

- Alreck, P.L., & Settle, R.B. (1995). *The survey research handbook* (2nd ed.). Chicago: Irwin.
- Bauer, B.J. (2006). *A success paradigm for project managers in the aerospace industry*. Unpublished doctoral dissertation, Capella University, Minneapolis, MN.
- Berger, L.A., & Berger, D.R. (2003). *The talent management handbook: Creating organizational excellence*. New York: McGraw-Hill Professional.
- Berthal, P.R., Colteryahn, K., Davis, P., Naughton, J., Rothwell, W.J., & Wellins, R. (2004). *Mapping the future: New workplace learning and performing competencies*. Baltimore, MD: ASTD Press.
- Besselman, J., Arora, A., & Larkey, P. (2000). Buying in a businesslike fashion and paying more? *Public Administration Review*, 60, 421-434.
- Borich, G.D. (1980). A needs assessment model for conducting follow-up studies. *Journal of Teacher Education*, 31(3), 39-42.
- Crawford, L. (1997). A global approach to project management competence. In *Proceedings of the 1997 AIPM National Conference* (pp. 220-228). Brisbane: Australian Institute of Project Management.
- Crawford, L. (2006). Developing organizational project management capability: Theory and practice. *Project Management Journal*, 37(3), 74-86.
- Cyert, R.M., & March, J.G. (1992). *A behavioral theory of the firm* (2nd ed.). Cambridge, MA: Blackwell Business.
- Draganidis, F., & Mentzas, G. (2006). Competency based management: A review of systems and approaches. *Information Management & Computer Security*, 14(1), 51-64.
- Fox, J.R., & Miller, D.B. (2006). *Challenges in managing large projects*. Fort Belvoir, VA: Defense Acquisition University Press.
- Gadeken, O.C. (2004). The ideal program manager. *Defense AT&L*, 33(3), 2-4.
- Gale, A. (2007). Competencies: Organizational and personal. In P.W.G. Morris & J.K. Pinto (Eds.), *The Wiley guide to project organization & project management competencies* (pp. 143-167). Hoboken, NJ: Wiley.



- GAO. (2005). *Better support of weapon system program managers needed to improve outcomes* (GAO-06-110). Washington, DC: Author. Retrieved April 10, 2007, from <http://www.gao.gov>
- GAO. (2008). *Defense acquisitions: Results of annual assessment of DoD weapon programs* (GAO-08-674T). Washington, DC: Author. Retrieved July 4, 2008, from <http://www.gao.gov/new.items/d08674t.pdf>
- Garman, A.N., & Johnson, M.P. (2006). Leadership competencies: An introduction. *Journal of Healthcare Management*, 51(1), 13-17.
- Golob, M.P. (2002). *Implementing project management competencies in the workplace*. Unpublished doctoral dissertation, Capella University, Minneapolis, MN.
- Jehn, K.A., & Mannix, E.A. (2001). The dynamic nature of conflict: A longitudinal study of intragroup conflict and group performance. *Academy of Management Journal*, 44(2), 238-251.
- Jones, R.E. (1997). *Partnering between government and industry: An acquisition reform initiative*. Arlington, VA: Science Application International Corporation (SAIC). Retrieved July 20, 2008, from <http://handle.dtic.mil/100.2/ADA328981>
- Joseph, E.E., & Winston, B.E. (2005). A correlation of servant leadership, leader trust, and organizational trust. *Leadership & Organization Development Journal*, 26(1/2), 6.
- Kadish, R. (2006). *Defense acquisition performance assessment report*. Retrieved March 30, 2006, from <http://www.acc.dau.mil>
- McClelland, D.C. (1973). Testing for competence rather than for intelligence. *American Psychologist*, 28, 1-14.
- Nissen, M.E., Snider, K.F., & Lewis, I.A. (2002). United States defense acquisition research program. In D.P. Slevin, J.K. Pinto & D.I. Cleland (Eds.), *The frontiers of project management research* (pp. xii). Newtown Square, PA: Project Management Institute.
- Parry, S.B. (1998). Just what is a competency? (And why should you care?). *Training*, 35(6), 58-64.
- Pinkerton, W.J. (2003). *Project management: Achieving project bottom-line succe\$\$*. New York: McGraw-Hill.
- Project Management Institute (PMI). (2002). *Project manager competency development framework*. Newtown Square, PA: Author.
- Ruth, D. (2006). Frameworks of managerial competence: Limits, problems, and suggestions. *Journal of European Industrial Training*, 30, 206-226.
- Samuel, T. (2003, June 23). The Hawkeye on the hill. *US News & World Report*, 134(22), 22. Retrieved February 13, 2009, from Business Source Complete database.
- Sue, V.M., & Ritter, L.A. (2007). *Conducting online surveys*. Thousand Oaks, CA: Sage Publications.
- Taylor, F.W. (1911). *The principles of scientific management*. New York: Harper & Brothers.
- Wells, C.V., & Kipnis, D. (2001). Trust, dependency, and control in the contemporary organization. *Journal of Business and Psychology*, 15(4), 593-603.
- Whiddett, S., & Hollyforde, S. (2003). *A practical guide to competencies: How to enhance individual and organisational performance* (2nd ed.). London: Chartered Institute of Personnel and Development.
- Womack, J.P., Jones, D.T., & Roos, D. (1991). *The machine that changed the world: How Japan's secret weapon in the global auto wars will revolutionize Western industry*. New York: HarperPerennial.



Panel #25 – Acquisition Management in a System-of-Systems World

Thursday, May 13, 2010

3:30 p.m. –
5:00 p.m.

Chair: Rear Admiral William E. Landay, III, US Navy, Program Executive Officer SHIPS

Contracting for Complex Products

David Van Slyke, Syracuse University; Trevor Brown, Ohio State University; Matthew Potoski, Iowa State University

Research on Systems-of-Systems Acquisition

Rene Rendon, Thomas Huynh and **John Osmundson**, Naval Postgraduate School

Creation of a System of Systems Portfolio Management and Technology Selection Methodology

Richard Volkert and Carly Jackson, SPAWAR Systems Center-Pacific; Peter Gentile, Lance Harper, Art VanNostrand, and Tom Sondi, Northrop Grumman Corporation; Kenneth Michaud, Pioneering Evolution



Contracting for Complex Products

David Van Slyke

Introduction

The US Federal Government spends just under twenty percent of its budget buying everything from paper clips to complex weapons systems. Effective contracting promises win-win exchanges: governments gain efficiency and qualities not available through in-house production, and vendors win because the price is above their production costs. Markets are most likely to produce win-win outcomes when buyers and sellers can easily define and verify product cost, quality and quantities. We call these simple products. Markets for simple products tend to have large numbers of buyers and sellers who are well informed about each others' offerings, can easily enter and exit the market, and can clearly define the terms of exchange. In such ideal circumstances, contracts are relatively complete in that there are few unanticipated circumstances in which the buyers' and sellers' roles are not clearly defined. If for some reason a buyer or seller fails to live up to her obligations, the transgression is quickly and easily recognized and a richly competitive market provides a replacement partner seeking similar terms.

When markets fail, the win-win outcomes of contracting are replaced by lose-lose or win-lose outcomes where the winner's gains are greater than the loser's losses. One source of market failure is buyer and seller uncertainty about the product in the exchange, what we call complex products.⁸³ Unlike simple products, the cost, quality and quantity parameters of complex products can not be easily defined or verified, leaving buyers and sellers unable to clearly and completely define exchange terms (Bajari & Tadelis, 2001).⁸⁴ The risk is that the government is the only purchaser and once the contract is let, the vendor is the only viable supplier, leaving each with no easy exit from the contract, limited information about costs and quality, and engaging a partner relatively unconstrained by market pressures. The consequence is a collective action problem in which the buyer and

⁸³ In other cases, either the buyer or seller may be more uncertain about the product. In cases where the seller has an information advantage, buyers cannot discern between high- and low-quality products and consequently pay only the price for low-quality goods less they be caught paying high-quality prices for low-quality goods. The result is a lemons market (Akerloff, 1970) in which the presence of low-quality products destroys the market for high-quality products. Cases where the buyer has an information advantage are less problematic: the buyer has every incentive to share information (stimulating vendor competition) and markets are efficient at distributing information.

⁸⁴ There are many sources of market failures, including incomplete property rights, transaction costs, information asymmetries, and barriers to market entry and exit (e.g., Mankiw, et al., 2002). Goods may be non-rivalrous or non-excludable so that transferable property rights cannot be established and enforced without transaction costs swamping gains from trade (Weimer & Vining, 1999). Historical accident may inefficiently lock-in path dependent technologies such as the QWERTY keyboard (David, 1985). A common thread in these cases is that the market failure is caused by the transaction costs stemming from limited information among participants, particularly buyers, or from goal incongruence between the buyers and sellers. Complex products can be viewed as just one type of market failure, albeit an extreme and difficult case.



seller have incentives to exploit contract ambiguities for their own gain at the other's expense, risking mutually disadvantageous outcomes.

In the business arena, the risk of market failure often justifies avoiding the market altogether, perhaps through vertically integrating production of the complex product (Williamson, 1991). In the public sphere, legal mandates often require governments to provide goods and services where markets are prone to fail and government production is impractical. Achieving public value in complex contracting requires transforming lose-lose conflict into win-win cooperation, a challenging though not impossible task, as the voluminous collective action literature suggests (e.g., Ostrom, 2000).

Our theoretical approach suggests that contracting relationships between buyers and sellers are more fruitfully specified along several dimensions. First, what is being exchanged in the relationship and how uncertain are the actors about the terms of exchange (e.g., are the products simple or complex?). Second, what aspects of the relationship are formally detailed (such as in a contract) and which are left informal (as in a network)? Third, how does the strategic context shape interactions between buyers (principals) and sellers (agents)? Self interest leads to win-win outcomes in contracts for simple products, for example, while complex products can become prisoners' dilemmas, as we show in this paper. Depending on the nature of the strategic context, other forms are possible.

In this paper, we first lay out the theoretical case for how complex contracting risks a collective action problem. Casting complex contracting as a prisoners' dilemma suggests potential avenues for achieving win-win outcomes, but it also reveals how uncertainty threatens cooperation and accountability mechanisms. Second, we illustrate the analytic value of the complex contracting theory with the case of the Coast Guard's controversial Deepwater project, a major acquisition program to upgrade and integrate its entire fleet of air and sea assets. As we show in this paper, the exchange between the Coast Guard and the private consortium which took the lead on producing and delivering Deepwater assets occurred in a highly uncertain environment, with an incomplete contract with both formally specified and informal elements, and in a prisoners' dilemma strategic context.

Complex Products and Complex Contracting

When making purchases, a buyer may be able to specify the objectives she wants to accomplish, but not the products and features which can accomplish them. In some circumstances, a buyer may be able to specify the objectives she wants to accomplish through purchasing, but not the products and features which can accomplish them. While complementary products can inform the buyer about product qualities, quantities and prices, sometimes the objectives are sufficiently unique that market signals about these products provide little guidance. For complex products, the buyer has value uncertainty—she does not know the value of different products' capabilities, qualities and tradeoffs among them—and the seller has cost uncertainty—he does not know the costs of producing the product with different capabilities and qualities to meet the buyer's objectives (Hart & Moore, 2008).

Incomplete Contracts

Contracts specify each party's obligations in an exchange, including the price, qualities, and quantities of the product. The buyer's obligations can include the payment terms and the terms under which the product is to be received (e.g. timing of delivery). The seller's obligations vary under different price arrangements, but generally include some



combination of output specifications such as product qualities and quantities and input characteristics such as time and materials.⁸⁵ The contract may also define each party's discretion, perhaps through reference to public law for default rights and obligations (Brown, Potoski & Van Slyke 2006).

Contract completeness is the degree to which the contract defines buyers and sellers' rights and obligations across all future contingencies (Hart & Moore 2008; Tirole, 1999; Bajari & Tadelis, 2001; Heinrich, 1999; Martin, 2004; O'Looney, 1998). All contracts are incomplete to some degree because the future contains an infinite number of scenarios, not all of which can be specified in advance. At some point the costs of writing contract terms for all future scenarios exceed the mutual gains from the trade, and no contracting would occur. Because of ambiguity about how to design and build a product to meet the buyer's objectives, and the associated costs faced by the seller, the two parties cannot specify all contract elements in advance. Complex products lead to highly incomplete contracts.

Asset Specific Investments

There are two primary means to reduce complex products' value and cost uncertainty – research and development and producing the product (i.e., learning by doing). Buyers can contract for both designing and/or building from sellers. Negotiating and executing a contract for a complex product often requires buyer and seller to make asset specific investments. Expenditures are asset specific to the extent they have no economic value outside the product being produced (Williamson, 2005). For example, some research in the US space program produced economic value outside the contract (e.g., Tang), while other research produced little value outside the contract (e.g., spacesuits, at least as of 2008). Other asset specific investments in complex contracting include the buyer and seller customizing production processes to suit each others' idiosyncrasies. Asset specific investments are lost if the contract is not executed.

Lock-In Risks

The consequence of an incomplete contract, asset specific investments and an unpredictable future is the classic “lock-in” problem (Williamson, 1996). A party becomes locked into a contract because it cannot redeploy its asset specific investments to other profitable endeavors; the other party can exploit unforeseen events and contract ambiguities. For the buyer, the “lock-in” risk is that once a seller has been selected, no other potential sellers have made the necessary investments, and the advantaged seller may look to opportunistically exploit contract ambiguities perhaps by “gold plating” the product with costly features that increase his profits, but for the buyer add little value and considerable expense. Likewise, because the seller has only one buyer for its products, the buyer may also opportunistically exploit contract terms for its own favor. The buyer may force a seller, for example, to make changes to a product that raise costs above the negotiated price even though it well knows that a cheaper product would meet her needs almost as well. In these circumstances, the exploiter's gains can be smaller than the exploited party's losses.

⁸⁵ Fixed-price contracts set compensation on the seller's outputs while cost reimbursement contracts set compensation on inputs, such as time and materials. These terms specify who generally bears cost risk: fixed price contracts place more of the cost risk on sellers and cost plus contracts place more of the risk on buyers (Bajari & Tadelis, 1999).

Absent lock-in problems, the buyer can simply replace an opportunistically behaving seller with a more suitable one, and a seller can likewise replace an opportunistic buyer. Lock-in problems coupled with incomplete contracts weaken the disciplining power of market forces. Within a single complex contract, each party is likely to find itself simultaneously advantaged in some areas and vulnerable to exploitation in others. Lock-in problems transform a contract from a market exchange to a political relationship whose outcomes are determined less by market forces and more by the strategic relationship between the buyer and seller.

Strategy

When performing in areas where the contract is vague, the buyer and seller's behavior can be either *perfunctory* or *consummate*, using the terminology of Hart and Moore (2008) and Williamson (1975, p. 69). Perfunctory behavior conforms to the bare minimum "letter" of the contract as enforceable by a court of law, while consummate behavior goes beyond what the bare minimum of the contract requires and towards greater win-win gain. Perfunctory behavior means accepting greater individual gain, but an almost certainly smaller mutual payoff. Consummate behavior means forgoing a large unilateral gain in exchange for a smaller individual payoff from a potentially larger mutual gain. For example, the complete portion of a contract can specify the number of jokes a hired comedian must tell, but it is impractical to specify how funny she should be when she tells them. A consummate comedian would strive for big laughs while a perfunctory comedian would settle for mild giggles. The degree to which the payoffs and penalties from consummate and perfunctory behavior affect contract behavior increases the more the contract is incomplete and the greater the lock-in problems. The goal of contract management is to develop and foster consummate behavior.

Complex Contracting as a Prisoners' Dilemma

The strategic implications of the lock-in contracting problem resemble a prisoners' dilemma. While not all contracting relations end up as prisoners' dilemma problems—buying simple products, for example, is likely to produce win-win outcomes—viewing complex contracting through the prisoners' dilemma (PD) lens helps diagnose the problem – why the failure risk is so high for these contracts—and identify its solutions—how to manage these contracts to transform lose-lose conflict into win-win cooperation. In a prisoners' dilemma game, two (or more) players choose whether to cooperate or defect, with the payoff of their choices jointly determined. If both choose to cooperate, each receives a moderately high payoff and if both choose to defect, each receives a moderately lower payoff. If one elects to cooperate while the other defects, the defector receives a very high payoff and the cooperator a very low one.

The buyer's and seller's complex contracting strategy options are analogous to prisoners' dilemma strategies (see Table 1). Perfunctory performance is analogous to defecting: it increases the performer's payoff but by a smaller amount than the reduction in payoff for the other party. Consummate behavior produces the higher mutual payoff for both, but risks the suckers payoff (a payoff of 1 in Table 1) if the other player elects to pursue perfunctory performance. The size of the payoffs in Table 1 increases with the degree the contract is incomplete and the degree of asset specific lock-in problems. For example, perfunctory seller behavior causes little harm when buying a simple commodity, like flour, and absent lock-in problems, troublesome sellers can easily be replaced with agreeable ones. Our examples assume symmetrical payoffs for buyers and sellers. In



practice, one side's advantages are likely to be greater, although the underlying logic of our theory holds so long as each side has some lock-in advantage over the other.

Table 1. The Complex Contract Dilemma

		<i>Buyer</i>	
		Consummate	Perfunctory
<i>Seller</i>	Consummate	3, 3 S, B	1, 4 S, B
	Perfunctory	4, 1 S, B	2, 2 S, B

Achieving win-win complex contracting outcomes requires changing the buyer's and seller's incentives from perfunctory behavior (defection) to consummate (cooperation) performance. The challenge therefore is to change the payoff structure to another type of game (Lichbach, 1996). First, the buyer and seller can seek credible commitments to cooperate, such as by submitting themselves to external supervision. For example, perfunctory behavior may damage a firm's reputation, making it a less attractive partner for future complex contracts. Second, if both the buyer and seller care about the future and the PD game is played over multiple rounds, cooperation can be achieved through a "tit-for-tat strategy" in which both sides initially cooperate and then mirror the other party's behavior from the previous round.

The Impact of Nature

While changing the payoff structure of the game increases the possibility of cooperation for complex contracting, external uncertainty, or ambiguity about the impact of future events or states (Heide & Miner, 1992), can undermine the prospects of cooperation, particularly in repeated play. Summarizing the complex contracting game as a decision tree illustrates how the parties can achieve cooperation, but also how the uncertainty inherent in complex contracting can make cooperation more difficult. Figure 1 depicts graphically the complex contracting game in tree form, with payoffs for cooperation and defection carried over from Figure 1, so that the payoffs for mutual cooperation are three each, for mutual defection two each, and in case of asymmetric strategies, one for the cooperator and four for the defector.

Phase I: Negotiation	Phase II: Strategy Selection		Phase III: Nature	Phase IV: Payoffs		Phase IV: Payoffs
	<i>Buyer</i>	<i>Seller</i>		<i>Buyer</i>	<i>Seller</i>	<i>Net Gain</i>
Incomplete Contract	Cooperate	Cooperate	Positive	$3 + 2 = 5$	$3 + 2 = 5$	10
			Negative	$3 + 0 = 3$	$3 + 0 = 3$	6
		Defect	Positive	$1 + 2 = 3$	$4 + 2 = 6$	9
			Negative	$1 + 0 = 1$	$4 + 0 = 4$	5
	Defect	Cooperate	Positive	$4 + 2 = 6$	$1 + 2 = 3$	9
			Negative	$4 + 0 = 2$	$1 + 0 = 1$	5
		Defect	Positive	$2 + 2 = 4$	$2 + 2 = 4$	8
			Negative	$2 + 0 = 2$	$2 + 0 = 2$	4

Figure 1. The Complex Contracting Game

The tree form represents complex contracting in stages. In the first stage, the buyer and seller negotiate contract terms, leaving some portion incomplete, and each makes asset specific investments. In the second phase, the players choose whether to cooperate or defect. Next, nature changes circumstances. As the contract is executed and the product is researched, developed and produced, nature reduces cost and value uncertainty by producing circumstances that affect the payoffs buyers and sellers receive. However, buyers and sellers do not necessarily know how nature affected them or the other party: some portion of these changed circumstances is known to both parties and others are known only to one party or the other. Perhaps the product was happily cheaper to produce for the seller and more valuable to the buyer than anticipated. Or, perhaps the seller's production costs were much higher and the product turned out to be less valuable to the buyer.

Payoffs

In the last stage, the payoffs are revealed to the buyer and seller, as determined by a combination of the parties' strategies and the effect of nature. For illustrative purposes, we adopt simplifying assumptions about the payoff distribution. The complex contracting decision tree and payoffs can of course be expanded to include asymmetrical effects of nature and payoffs for cooperation and defection. Figure 2 implicitly assumes nature equally advantaged and disadvantaged both the buyer and seller, with a plus two in case of favorable circumstances and zero in the case of unfavorable circumstances. A central lesson revealed in the tree form of the game is that the payoffs that buyers and sellers

receive can depend as much on external circumstances as on their strategic choices about cooperation. In Figure 1, for example, the seller receives the same payoff (three) under mutual cooperation with unfavorable circumstances as she receives under a suckers payoff (seller cooperates, buyer defects) with favorable circumstances. If the buyer and seller do not know either the other party's strategy or the whether circumstances were favorable, neither will know the extent to which the payoff was due to their strategic choice or the vagaries of natural fortune. Moreover, without credible verification, neither side can reliably claim it was cooperating.

Complex Contracting and the Coast Guard's Deepwater Program

In this section we present the Coast Guard's Deepwater program to illustrate the challenges inherent in contracting for complex contracts. The Deepwater program is an effort to upgrade and overhaul the Coast Guard's "deepwater" air and sea vessels and the command and control links among them.⁸⁶ In 1998, the Coast Guard issued a Request for Proposal to design a system of interoperable air and sea assets to meet its mission and performance goals while lowering total ownership costs. The Coast Guard evaluated three industry proposals and selected a system design from Integrated Coast Guard Systems (ICGS, a consortium between Lockheed Martin and Northrop Grumman).⁸⁷ Under the ICGS proposal, the Coast Guard's Deepwater assets would all be fully integrated in a state-of-the-art command, control, communications, computers and intelligence, surveillance, and reconnaissance system, commonly referred to as C4ISR. In June 2002, the Coast Guard awarded its first Deepwater contract to ICGS.

The Deepwater case summary and analysis presented below are based on extensive qualitative research. Our first data source is a thorough review of the vast public record on the program. The General Accounting Office has written over a dozen reports on Deepwater (e.g., GAO, 2004, March; 2005, April; 2005, July; 2007, February; 2007, March; 2007, June; 2008, March; 2008, June) and the Inspector General of the Department of Homeland Security reported on a major internal Deepwater investigation (Department of Homeland Security Office of Inspector General, 2007). Deepwater has been subject to several House and Senate committees hearings,⁸⁸ Congressional Research Service Reports⁸⁹ and other Deepwater reports from the Defense Acquisition University (2007) and Acquisition Solutions (2001). The Deepwater program has also been reported fairly extensively in the news media, although we do not rely much on media reports for our primary case data.

⁸⁶ The term "deepwater" refers to Coast Guard assets that operate in literal deepwater, 50 miles off shore.

⁸⁷ ICGS' proposal included five new sea vessels, two fixed-wing aircraft, two helicopters, and one unmanned aerial vehicle. The other assets were upgrades.

⁸⁸ For examples, see the Committee on Homeland Security, Subcommittee on Border, Maritime, and Global Terrorism, <http://homeland.house.gov/Hearings/index.asp?ID=50> , and the Subcommittee on Coast Guard and Maritime Transportation at <http://transportation.house.gov/Media/File/Coast%20Guard/20070130/Opening20070130.pdf>

⁸⁹ See, for example, <http://fas.org/sgp/crs/weapons/RS21019.pdf> and <http://www.fas.org/sgp/crs/weapons/RL33753.pdf>



We have also conducted more than one hundred interviews with individuals knowledgeable about the case, including former and current Congressional committee staff members, current and former Coast Guard officials, officials from ICGS, officials from industry not associated with ICGS, officials from the Department of Homeland Security, academic experts on federal acquisition, and officials who serve in an advisory capacity to members of Congress. A semi-structured interview protocol was used and all interviews were conducted by two, and in some cases three, researchers. The interview participants were guaranteed confidentiality and anonymity, and participation was completely voluntary with the option of withdrawing at any point. The project was approved by the institutional review boards at each of the author's institutions. Interview notes were transcribed and coded, and the findings triangulated against other data sources.

In the remainder of this section, we place the Deepwater contract in the complex contracting theoretical perspective. Deepwater negotiations occurred in an environment of high uncertainty and lock-in fears that produced an incomplete contract. The Coast Guard and ICGS both claimed to choose cooperative, consummate contracting strategies, although these claims are difficult to verify. The Coast Guard and ICGS received lower than anticipated payoffs, however it is not clear whether this was the result of their contract strategies or an unexpectedly difficult contracting environment. Our conclusions about the Deepwater processes and outcomes are drawn from our interviews and systematic analysis of the publicly available Deepwater information, and those documents received through the Freedom of Information process or provided by the interviewees.

Background

In recent history the Coast Guard's procurement practice has been to separate purchases for individual classes of assets—ships, cutters, planes and helicopters; when a class of ships was no longer sea worthy, the Coast Guard bought a new one to replace it, perhaps with a modified design better suited to the Coast Guard's evolving mission. Because it bought fewer and smaller assets relative to other major naval buyers—notably the US Navy—the Coast Guard largely made ad hoc purchases from a handful of small sellers (e.g., regional shipyards such as Bollinger). By the early 1990s, it became clear that the Coast Guard needed a more targeted and strategic approach to upgrade its rapidly aging assets.⁹⁰ Many Coast Guard assets were reaching the end of their usable life-span and were ill suited to the modern Coast Guard's missions. The Coast Guard leadership lobbied Congress for a long-term acquisition strategy that would upgrade and modernize a significant portion of the Coast Guard fleet with a stable funding stream. In 1998, Congress and the Clinton administration committed to a multi-year procurement at approximately \$500 million a year, significantly more than the Coast Guard's historical acquisition expenditures.⁹¹ The result was the Deepwater program or Project Deepwater.

Deepwater as a Complex Product

The Deepwater program was a complex product. The Coast Guard was highly uncertain about the value it would receive from Deepwater products. The Coast Guard

⁹⁰ As of 2001, 86% of the Coast Guard's water and air assets had reached or were expected to reach the end of their planned service life within five years. The Coast Guard's fleet of assets was widely considered to be one of the oldest in the world, ranking 37 out of 39 of the fleets worldwide (Acquisition Solutions, 2001, p. 6).

⁹¹ <http://govinfo.library.unt.edu/npr/library/news/062999.html>



understood its objectives—maritime security (upholding the law), maritime safety (rescuing the distressed), protection of natural resources (protecting the environment), maritime mobility (ensuring safe marine transportation), and national defense (operating in coordination with the US Navy)—but it lacked information about the options for achieving them. The Coast Guard knew the basic components that would ultimately comprise its asset fleet—small and large boats, planes, and helicopters, tied together through communication and integration technologies. But the Coast Guard did not know what each of their performance specifications would be and how they would operate together in a system. As a result, the Coast Guard was not well positioned to assess alternative system portfolios and system components. For example, how many fewer aircraft would be needed if the performance of the large cutters were increased twenty percent? ICGS had cost uncertainty. The Coast Guard had established a hard overall-cost cap (\$500 million annually), but ICGS did not know how much it would cost to deliver these to meet the Coast Guard’s objectives. Full cost information for each asset would not be available until the Coast Guard either specified performance standards with some precision or authorized a first-in-class design.⁹² Any delay in identifying these specifications would compound costs as production processes lay idle.

Asset Specific Investments and Lock-In Risks

The Deepwater assets vary in the degree to which they required asset specific investments. On the low end of asset specificity are alterations to existing, highly marketable assets. For example, ICGS can turn to a variety of subcontractors to provide upgrades to helicopter engines. On the high end of asset specificity are new assets for which the Coast Guard is one of only a few potential buyers, if not the only buyer. For example, one of the primary assets of the Deepwater program is the National Security Cutter (NSC), the largest class of ships in the Coast Guard fleet. Northrop Grumman—the NSC lead contractor—developed a specialized production process for this asset because of the Coast Guard’s unique performance needs. The Coast Guard also made investments that were to a degree asset specific, notably the creation of a Deepwater acquisition office, separate from its existing procurement infrastructure. This unit was to design procurement practices and systems exclusively for ICGS; the Coast Guard staff needed special clearance to work on the Deepwater acquisition. As processes and staff become embedded with a single seller, it becomes harder to adapt to others.

For both parties, these asset specific investments created lock-in risks. For the Coast Guard the risk was that ICGS would gain an information advantage as it designed and built the assets. If ICGS elected to abuse its information advantage (e.g., by “gold plating” the product), the Coast Guard would have limited options for alternatives. For example, a highly asset specific element of ICGS’ Deepwater system is the logistics and communication systems for integrating all the system components, known as C4ISR.⁹³ ICGS has

⁹² First-in-class designs typically encounter cost overruns and schedule delays as the buyer and seller work out the precise specifications for the product. Cost-plus contracts are often used for the prototype design where the buyer bears greater cost risk. This puts the burden on the buyer to determine what it is exactly that they want to buy. Once the first-in-class asset is completed, parties often switch to fixed-price contracts for subsequent assets. In this case the seller bears greater cost risk.

⁹³ Lockheed Martin assumed the lead in development and production of this element because of its experience with the Navy’s AEGIS system (http://en.wikipedia.org/wiki/Aegis_combat_system). This



substantial research, development, modification and adaption costs invested in tailoring this system to the Coast Guard's needs. In order to secure a return on its investment, ICGS has a strong incentive to make this system as proprietary as possible. As such, the Coast Guard now has to largely rely on ICGS for costly capability enhancements, training, technical assistance, and standard operational maintenance and upgrades. For ICGS, the primary lock-in risk is that it may not recoup its investments (i.e. research and development and production) if the Coast Guard elects to stop buying those system elements for which ICGS has made asset specific investments. If the Coast Guard elected to buy helicopter engine upgrades from an alternative supplier, ICGS can tender its engines to another buyer—the market is thick. The risk is for highly asset-specific elements like the C4ISR system developed for the Coast Guard.⁹⁴

Incomplete Contracts

One can imagine that in designing and producing an aerial or sea vessel, let alone an interoperable system of such assets, there is an almost infinite set of product specifications over which the buyer and seller might negotiate (e.g., the speed and lift of helicopters, the time at sea for boats, crew capacity, etc.). To fully specify many of these decisions requires forecasting an array of variable future conditions (i.e., states of nature), some of which are highly unpredictable (e.g., weather, terrorists, drug runners). Given the value uncertainty faced by the Coast Guard, the cost uncertainty faced by ICGS, and the mutual risk of lock-in, the two parties entered into a contract arrangement that specified some aspects of the system, but left others unspecified.

The Coast Guard and ICGS sought to balance their need for specificity against the uncertainty inherent in buying a complex product by specifying three contract layers. The top layer was a performance-based indefinite delivery, indefinite quantity (IDIQ) contract. In the simplest terms, an IDIQ does not specify a firm quantity of products or the tasks required to produce them, but rather specifies a minimum or maximum number of products and some end-point for termination of the agreement.⁹⁵ The Deepwater IDIQ contract specified that the Coast Guard could buy a set of system components without competitively bidding each one; instead each purchase was only with ICGS. In a sense, the IDIQ acts like a menu with the base costs for various items, but where the add-ons are neither specified nor priced.

The middle layer of the contract set the broad terms of exchange. Each purchase under the IDIQ was negotiated through a task order between the Coast Guard and ICGS that specified basic terms, such as the number of units to be purchased in a class of assets and their delivery schedule, but left many dimensions indeterminate, like the exact design and performance specifications, cost schedules, and the evaluation metrics.

Specifying many of the details of each task order occurred through a final contract layer that was intended to facilitate less formalized cooperation through the process of designing, testing, and building the asset. This process occurred through Integrated Product Teams (IPTs). The IPTs were designed as collaborative governance mechanisms

is a legacy system which is the first generation predecessor to a more modern Coast Guard C4ISR system.

⁹⁴ Interestingly, there is significant interest from overseas buyers for the C4ISR system, but because of national security concerns ICGS is prohibited from selling comparable systems abroad.

⁹⁵ See e.g., *FAR Subpart 16.5—Indefinite-Delivery Contracts* (http://www.arnet.gov/far/current/html/Subpart%2016_5.html#wp1093133)



that brought together ICGS personnel, subcontractors, and Coast Guard officials to decide the important details about the asset under their jurisdiction. However, once production was underway, rather than fully renegotiate each task order to reflect some design change or refinement, the Coast Guard and ICGS used Undefined Contract Actions (UCAs). UCAs are a legal vehicle that allows production to continue after a design change, even though the parties had not formally negotiated the full price and terms of that change. UCAs require that the parties formally resolve the specification and price within 180 days. UCAs place the cost risk on the buyer because the seller has considerable discretion over the price charged for each revision. Once these items become definitized they operate like fixed cost contracts.

These three contractual layers specified the terms of the exchange and established a process for the Coast Guard and ICGS to reduce value and cost uncertainty, although they still did not fully specify the contract. The task orders and UCAs could not specify every design and performance requirement for each asset and how they would fit together. Part of the specification challenge was the interoperability requirement—and hence interconnectedness—of the system. Any performance specification (e.g., the speed of a boat) had implications for other system elements (e.g., the range of helicopters and planes). Because the acquisition of the total Deepwater fleet was sequenced over a twenty-five year period, the Coast Guard and ICGS were unable to forecast (and specify formally in a task order) every detail of the assets early in the acquisition because these would then cement performance specifications for all later assets, which had not yet been specified or fully designed in some cases. Furthermore, because so much about these assets would not be known until the first unit was built, tested and refined in operating conditions, experience was perhaps the best guide for specifying much of the performance terms. Many of these elements would need to be resolved either informally by ICGS and the Coast Guard working together or by one or the other of the parties deciding unilaterally. As a result, the incompleteness of the contractual arrangements left substantial room for the behavior, or strategy, of both the Coast Guard and ICGS to impact outcomes.

Strategy

At the outset of the contract, ICGS and Coast Guard pledged to pursue consummate, cooperative strategies for how they treated the contract, often invoking “partnership” language to describe the relationship they were entering. Harkening back to our theoretical terms, a partnership implies that each party was prepared to forgo an advantage for itself in favor of a larger benefit for the other side. Perfunctory behavior, on the other hand, means following the “letter” of the contract and perhaps engaging in opportunistic behavior to pursue unilateral gain at the expense of greater mutual gain. With so much of the Deepwater contract incomplete, much of the contract’s payoff would be determined by the parties’ behavior.

Here we focus on one of the first task orders—the delivery of the NSC—to illustrate the potential impact of strategy on outcomes. The NSC presented lock-in risks given the high degree of asset specificity described earlier. The task order was incomplete in that it left many performance requirements indeterminate, following the pattern of devolving these decisions to an IPT and later definitization.⁹⁶ In addition, the task order was incomplete in

⁹⁶ The design of the NSC has occurred in two phases. In the Phase 1 RFP, 85% of the design criteria and performance standards had been developed by the Coast Guard and the American Bureau of Shipping. In the second phase, the contract, ICGS had discretion over what the remaining criteria and standards were going to be. However, the Coast Guard did not include a contractual mechanism that

that it did not identify the decision making rights and obligations of either ICGS or the Coast Guard over these unspecified elements. Specifically, the task order did not identify: which party had decision making authority over structural design specifications; the conditions under which independent third-party assessment of the design would be necessary, or which organizations would be qualified to perform this role (e.g., US Navy's Surface Warfare Division); corrective action or sanctions in the event that design specifications were not certifiable; criteria and evaluation process for paying award fees; and penalties or accommodations for cost overruns or missed delivery dates. This left both parties with a wide berth to behave consummately or perfunctorily.

In these areas, ICGS had discretion to openly discuss and jointly agree with the Coast Guard on the remaining unspecified decisions for the final design of the NSC. Such an approach would be consummate behavior because it would produce win-win outcomes for the Coast Guard and ICGS. Alternatively, the contract allowed ICGS to specify design standards unilaterally). Such an approach would be perfunctory behavior because, although it was easier for ICGS to decide standards on its own, Coast Guard bore the risk of buying assets whose performance abilities did not meet what it needed. If ICGS chose consummate behavior, it would use the IPT to jointly discuss and develop design and performance standards that aligned with the Coast Guard's goals for the cutter (e.g., that it be capable of being at sea for 230 days a year with a 30-year service life). ICGS would explain to the Coast Guard the value and cost tradeoffs associated with different design and performance standard modifications.⁹⁷ Placing this discussion within the IPT might also provide ICGS with the perspectives of both Coast Guard technical and operational experts along with third party experts, such as those from the American Bureau of Shipping.

This cooperative approach through the IPTs would increase ICGS's costs because orchestrating an open conversation with the Coast Guard about the comparative pros and cons of alternative design specifications would require significant investment of staff and time.⁹⁸ There would also be no guarantee that the investment would benefit ICGS more than these costs, in large part because ICGS would have to keep production processes idle while it waited for Coast Guard officials to decide each unspecified item. However, supplying more complete information would provide more value to the Coast Guard than it would cost ICGS to perform because the Coast Guard would be more likely to receive an asset—the NSC in this case—better aligned with mission requirements. Conversely, ICGS could spend little on providing information to the Coast Guard, or simply make decisions itself, which would be less costly. However, because the Coast Guard would receive a product that might not meet its performance requirements, it would then be forced to pay even more to modify the asset. Such changes would likely cost more for the Coast Guard than they could have cost ICGS (and the Coast Guard) to negotiate in advance through the IPTs if ICGS had been pursuing a consummate approach to the contract.

would ensure that the alternative standards would be consistent with the standards developed in the Phase 1 RFP. (http://www.dhs.gov/xoig/assets/OIGtm_RLS_051707.pdf).

⁹⁷ For example, these discussions might include an explanation on the part of ICGS would clearly articulate that an additional investment of X dollars would yield an increase in Y days at sea or conversely that accepting some alternative would lead to a Z percent decline in days at sea, but potentially lower costs by X dollars or increase the lifespan of the cutter by Y years.

⁹⁸ This would require conducting simulations and cost-benefit modeling to determine the relative pros and cons associated with each design alternative, coordinating participation from all the IPT participants, and deploying technical and contractual staff to the IPT to present the information.



Just as ICGS could pursue a consummate or perfunctory approach to the contract, the Coast Guard faced a similar decision. The Coast guard could reduce ICGS' costs by inviting ICGS to provide comparative information on alternative performance specifications and quickly responding with decisions. Such an approach would be consummate behavior. Alternatively, the Coast Guard could eschew opportunities to collaborate with ICGS to determine the design standards in the IPT, and instead unilaterally process UCAs, which would be perfunctory behavior. Choosing consummate behavior would mean the Coast Guard would actively participate in the IPT by suggesting ways the proposed assets might be designed or modified to fulfill mission requirements before production moved too far along. Such behavior would increase Coast Guard's costs because it would require collecting relevant performance requirements and translating it for ICGS. This action would produce less value for the Coast Guard because, in the absence of a contractual requirement that ICGS provide comparative information (recall that these obligations were not specified in the contract), ICGS may not respond to the request. However, such efforts to specify performance requirements would provide more value for ICGS because it would both produce a satisfied customer more inclined to renew existing and process new task orders and lower the costs of designing and producing subsequent interoperable assets.⁹⁹ Conversely, if the Coast Guard chose perfunctory behavior, it would save the cost of acquiring performance requirement information needed to specify incomplete design standards. However, ICGS's costs would increase by more than the Coast Guard's saving because ICGS would either have to devote its own resources to gather the information (which would likely cost more than if the Coast Guard did this itself) or increase the risk of the Coast Guard not renewing the contract.

Taken together, ICGS and Coast Guard found themselves on the horns of a collective- action problem in resolving the incomplete terms of the contract. Both the Coast Guard and ICGS would be the best off if each behaved consummately. The costs of cooperation—namely additional resources devoted to reducing the other party's value or cost uncertainty—would be high for each, but these costs would be less than the mutual gain—a highly satisfied Coast Guard enjoying a NSC that met its mission requirements and a likely continuation of the IDIQ contract and subsequent task orders for ICGS. However, the benefits of mutual cooperation were jeopardized by the risk of the other player's possible defection. The Coast Guard and ICGS risked losing their investments if the other party opted not to reciprocate (i.e., the sucker's payoff). If the Coast Guard behaved consummately and ICGS perfunctorily, the Coast Guard would receive an NSC that met basic contractual requirements (i.e., adhered to the "letter" of the contract) but failed to meet expectations in areas not specified in the contract, with the ICGS still receiving its full payment. If ICGS behaved consummately and the Coast Guard perfunctorily, ICGS would have spent effort to meet its best guess of the Coast Guard's mission requirements but may have still missed the mark, thereby subjecting itself to the risk of not receiving full payment and putting future task orders at risk. Finally, if both parties behaved perfunctorily by avoiding costly collaboration in the IPT, the result would be an NSC that failed to meet Coast Guard's needs and jeopardized future IDIQ and task orders.

⁹⁹ Once ICGS knows all the design specifications for the NSC, it knows many of the specifications for interoperable aspects of the other assets thus lowering its design costs over the long term.



Nature and Outcomes

On May 8, 2008, ICGS delivered the first-in-class NSC—the Bertholf—255 days¹⁰⁰ after the projected delivery date and over double the projected cost baseline.¹⁰¹ Preliminary testing of the Bertholf revealed 2800 issues (trial cards) to be addressed, but only eight issues (starred trial cards) which required addressing before acceptance.¹⁰² The boat proved sea worthy and met most contractually specified performance requirements; the Coast Guard accepted delivery.

Both sides received less than they anticipated and are generally unhappy with contract outcomes. Performance evaluation of both parties by each other and external parties and overseers has questioned a range of leadership actions, decisions, and cast blame on each for the cost overruns, quality deficiencies, and schedule delays (GAO, 2008, March; 2008, June; 2007, February; 2007, March; 2005, April; 2005, July; 2004; CRS, 2008; DHS OIG, 2007; DAU 2007). The Coast Guard faced program delays, higher costs, and was criticized for accepting a flawed asset design. ICGS saw Congressional scrutiny, reputational damage, and lost task orders under the initial IDIQ.

Our framing of the Deepwater contract as a prisoner's dilemma suggests a possible explanation for these outcomes is that both the Coast Guard and ICGS behaved perfunctorily. The Coast Guard has been criticized for not actively participating in the NSC IPT, ceding decision authority over many unspecified design elements to ICGS, and unilaterally changing some NSC specifications through UCAs and then failing to definitize (GAO, 2008, June). ICGS has been criticized for not providing the Coast Guard sufficient information about the NSC's value-cost tradeoffs, over-billing the Coast Guard for modifications and alterations, and in some cases improperly exercising independent decision authority in the IPT by specifying design standards without sufficient Coast Guard input (DHS OIG, 2007; GAO, 2008, June; 2007, February; 2007, March; 2007, June).

Recall that one of the implications of our complex contracting theory is that it is particularly difficult to sort out the independent impact of strategy and nature on outcomes. A negative turn of events outside the control of each party, for example, may make the outcomes of even a cooperative contract appear substandard. In the case of the NSC, unpredicted external events may have contributed to outcomes and thus masked what might have been cooperative, consummate contracting practices. First, the terrorist attacks of 9-11 (which hit in the midst of the NSC's design and production) spurred the assignment of Coast Guard to the newly created Department of Homeland Security and expanded its mission responsibilities. Coast Guard's enhanced mission prompted significant changes to the NSC, notably an increase in its size and capabilities. Given the speed of events and Coast Guard's desire not to derail the production of the NSC, Coast Guard made these decisions unilaterally. ICGS, perhaps interpreting the receipt of major changes to an agreed

¹⁰⁰ "The Deepwater contract originally called for production and deployment work for NSC1 to be completed on August 3, 2007 with final delivery to the Coast Guard scheduled for August 27, 2007" (page 7 of http://www.dhs.gov/xoig/assets/mgmttrpts/OIG_07-23_Jan07.pdf). The Coast Guard reports taking delivery of the first NSC on May 8, 2008 (<http://www.uscg.mil/acquisition/nsc/default.asp>). This represents a 255-day delay.

¹⁰¹ The original contract estimate for the NSC was valued at \$322.2 million. Modifications and alterations in the design lead to \$291.2 million in increased costs plus \$35 million for inflation adjustments (http://www.dhs.gov/xoig/assets/mgmttrpts/OIG_07-23_Jan07.pdf).

¹⁰² <http://www.piersystem.com/go/doc/786/201676/>



upon design as perfunctory behavior, did not invest in explaining the cost implications of these changes to Coast Guard. A second unforeseen event was Hurricane Katrina, which struck the Gulf Coast in 2005, damaging Northrop Grumman's Ingalls shipyard in Pascagoula, Mississippi where the NSCs were being constructed. The damage delayed delivery and increased production costs. As ICGS felt that the calamity was not its fault, it passed some cost increase to the Coast Guard. At the time of billing, ICGS submitted a Request for Equitable Adjustment (REA), which included modifications to meet the 9-11 mission changes and costs associated with Hurricane Katrina.

Although the exact impact of these unforeseen events remains unknown, Coast Guard and ICGS believe the other's perfunctory behavior is largely to blame for the NSC's outcomes. In the most recent task order contracts, Coast Guard and ICGS clearly adopted perfunctory strategies that meet the letter of the contract, without the spirit of cooperation. The pretense of collaboration has been abandoned as the Coast Guard has set up its own acquisition directorate, assumed an increasing number of the responsibilities delegated to ICGS under the original IDIQ contract, is exercising greater authority over decision making (in and out of the IPTs), and is seeking to buy assets outside the IDIQ with ICGS. The promise of a win-win "partnership" has deteriorated into a lose-lose transaction as each party pursues a perfunctory strategy.

Conclusion

The theory of complex contracting and the illustration of the Coast Guard's Project Deepwater presented suggest that complex contracting is risky. Complex contracts are highly uncertain, costly to negotiate and execute, and obfuscate accountability. Under the prisoners' dilemma, lose-lose defection is individually more attractive than win-win cooperation, despite cooperation's higher mutual gains. Uncertainty about nature's contributions to payoffs mean even the contract parties do not know whether contract outcomes stem from misfortune, the other party's malfeasance, or their own mismanagement. In the case of Deepwater, both Coast Guard and ICGS assumed the other "defected." Overseers, notably Congressional oversight committees, have pushed the Coast Guard towards a perfunctory contracting approach. As our analysis shows, however, while some evidence suggests that one or both of the exchange parties are partially at fault, nature may have produced negative outcomes and pushed the parties towards a lose-lose outcome. In moving forward, the default response should not necessarily be to position Coast Guard and whatever sellers it engages into a rigidly perfunctory posture.

Because complex contracts are prone to renegotiation, the "shadow of the future" opens a wealth of cooperative strategies to foster norms of reciprocity and trust and thus allow contract parties to turn lose-lose conflict into win-win cooperation (e.g., Axelrod, 1984; Heide & Miner, 1992). Moreover, these cooperative strategies enhance overseer's ability to hold the exchange parties accountable in the contract process. In the case of Deepwater, the multi-stage architecture of the contracting process allows for this kind of renegotiation. While the projection is that it will take 25 years to deliver all the component Deepwater assets, the Coast Guard, and ICGS did not formally commit to a 25-year agreement. Instead the IDIQ contract is structured in five-year increments. The asset-specific nature of the some of the system components means that the Coast Guard faces a thin market of alternative suppliers for some assets, but at a minimum the overall contract arrangement allows for both exit and renegotiation.

The shadow of the future shapes whether parties spiral away from partnership towards a more formalized and perhaps antagonistic relationship. Congress' increasing



oversight and involvement in the Deepwater program demonstrate how political actors can change the rules of the game. At present, Congress is pushing the Coast Guard towards a non-cooperative approach and away from a cooperative stance. Congress could instead foster cooperation, for example, rather than forcing the Coast Guard to move away from an integrated relationship with ICGS to a more differentiated purchasing arrangement, Congress could encourage the Coast Guard to build on recent efforts to enhance the Coast Guard's role in the IPTs and rely more extensively on third-party certification of product design and delivery. A primary justification for such an approach is that there are significant opportunities to capture knowledge and information from the first round and apply it to subsequent rounds of contracting. The practice of complex contracting need not be so dire.

References

- Acquisition Solutions. (2001, July 14). *Independent assessment of the United States Coast Guard integrated deepwater system acquisition*. Issue Brief. Arlington, VA: Author.
- Akerloff, G. (1970). The market for lemons: Qualitative uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3), 488.
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Bajari, P., & Tadelis, S. (2001). Incentives versus transaction costs: A theory of procurement contracts. *Rand Journal of Economics*, 32(3), 387-407.
- Bajari, P., & Tadelis, S. (1999, March 15). *Procurement contracts: Fixed price vs. cost plus*. Retrieved from <http://ssrn.com/abstract=156470>
- Brown, T.L., Potoski, M., & Van Slyke, D.M. (2006). Managing public service contracts: Aligning values, institutions, and markets. *Public Administration Review*, 66(3), 53-67.
- Cooper, P. (2003). *Governing by contract: Challenges and opportunities for public managers*. Washington DC: CQ Press.
- David, P.A. (1985). Clio and the economics of QWERTY. *The American Economic Review*, 75(2), 332-337.
- DAU. (2007, February). *Quick look study: United States Coast Guard Deepwater Program*. Department of Homeland Security, Office of Inspector General. (2007, January). *Acquisition of the national security cutter: US Coast Guard (OIG-07-23)*. Washington, DC: US GPO.
- GAO. (2004, March 9). *Contract management: Coast Guard's Deepwater program needs increased attention to management and contractor oversight* (GAO-04-380). Washington, DC: Author.
- GAO. (2005, April 20). *Coast Guard: Preliminary observations on the condition of Deepwater legacy assets and acquisition management challenges* (GAO-05-307T). Washington, DC: Author.
- GAO. (2005, July 22). *Coast Guard: Progress being made on addressing Deepwater legacy asset condition issues and program management, but acquisition challenges remain* (GAO-5-757). Washington, DC: Author.
- GAO. (2007, February 15). *Coast Guard: Preliminary observations on Deepwater program assets and management challenges* (GAO-07-446T). Washington, DC: Author.
- GAO. (2007, March 8). *Coast Guard: Status of efforts to improve deepwater program management an address operational challenges* (GAO-07-575T).
- GAO. (2007, June). *Coast Guard: Challenges affecting Deepwater asset deployment and management and efforts to address them* (GAO-07-874). Washington, DC: Author.
- GAO. (2008, March 5). *Coast Guard: Deepwater program management initiatives and key Homeland Security missions* (GAO-08-531T). Washington, DC: Author.



- GAO. (2008, June). *Coast Guard: Changes in course improves Deepwater management and oversight, but outcome still uncertain* (GAO-08-745). Washington, DC: Author.
- Hart, O., & Moore, J. (2008). Contracts as reference points. *The Quarterly Journal of Economics*, 123(1), 1-48.
- Heide, J., & Miner, A. (1992). The shadow of the future: Effects of anticipated interaction and frequency of contact on buyer-seller cooperation. *Academy of Management Journal*, 35(3), 265-291.
- Heinrich, C.J. (1999). Do government bureaucrats make effective use of performance management information? *Journal of Public Administration Research and Theory*, 9(3), 363-394.
- Kelman, S.J. (2002). Contracting. In S.M. Lester (Ed.), *The tools of government: A guide to the new governance*. New York, NY: Oxford University Press.
- Kettl, D. (1993). *Sharing power: Public governance and private markets*. Washington DC: Brookings Institution.
- Lichbach, M.I. (1996). *The cooperator's dilemma*. Ann Arbor, MI: University of Michigan Press.
- Martin, L.L. (2004). Performance-based contracting for human services: Does it work? *Administration in Social Work*, 29(1), 63-77.
- Milward, H.B., & Provan, K.G. (2000). Governing the hollow state. *Journal of Public Administration Research and Theory*, 10, 359-80.
- O'Looney, J.A. (1998). *Outsourcing state and local government services: Decision-making strategies and management methods*. Westport, CT: Quorum Books.
- O'Rourke, R. (2008, June 5). *Coast Guard Deepwater acquisition programs: Background, oversight issues, and options for Congress*. Congressional Research Service Report for Congress. Washington, DC: US GPO.
- Ostrom, E. (2000). Collective action and the evolution of social norms. *The Journal of Economic Perspectives*, 14(3), 137-158.
- Sclar, E.D. (2000). *You don't always get what you pay for: The economics of privatization*. Ithaca, NY: Cornell University Press.
- Smith, S.R., & Lipsky, M. (1993). *Non-profits for hire: The welfare state in the age of contracting*. Cambridge, MA: Harvard University Press.
- Tirole, Jean. (1999). Incomplete contracts: Where do we stand. *Econometrica*, 76(4): 741-781.
- Weimer, D., & Vining, A. (1999). *Policy analysis: Concepts and practice*. Upper Saddle River, NJ: Prentice Hall.
- Williamson, O. (1975). *Markets and hierarchies, analysis and antitrust implications: A study in the economics of internal organization*. New York: Free Press.
- Williamson, O. (1996). *The mechanisms of governance*. New York: Oxford University Press.
- Williamson, O. (2005). The economics of governance. *American Economic Association Paper and Proceedings*, 95(2), 1-18.



Research on Systems-of-Systems Acquisition

Rene Rendon—Rene G. Rendon is an associate professor at the Naval Postgraduate School, where he teaches defense acquisition courses. He served for over twenty years as a contracting officer in the USAF, retiring at the rank of lieutenant colonel. His career included assignments as a contracting officer for the Peacekeeper ICBM, Maverick Missile, and the F-22 Raptor. He was also the director of contracting for the Space Based Infrared satellite program and the Evolved Expendable Launch Vehicle rocket program. Rendon has published in the *Journal of Public Procurement*, the *Journal of Contract Management*, and the *Project Management Journal*.

Rene G. Rendon
Naval Postgraduate School
Monterey, CA 93943
rgrendon@nps.edu

Thomas V. Huynh—Thomas V. Huynh obtained simultaneously a BS in Chemical Engineering and a BA in Applied Mathematics from UC Berkeley and an MS and a PhD in Physics from UCLA. He is an associate professor of systems engineering at the Naval Postgraduate School in Monterey, CA. His research interests include uncertainty management in systems engineering, complex systems and complexity theory, system scaling, and system-of-systems engineering methodology. Prior to joining the Naval Postgraduate School in 2003, he was a Fellow at the Lockheed Martin Advanced Technology Center in Palo Alto and Sunnyvale, CA, where he engaged in research in computer network performance, computer timing control, bandwidth allocation, heuristic algorithms, nonlinear estimation, perturbation theory, differential equations, and optimization. While he spent 23 years in the aerospace industry, he was also teaching part-time in the departments of Physics and Mathematics at San Jose State University. Dr. Huynh is a member of INCOSE.

Thomas V. Huynh
Naval Postgraduate School
Monterey, CA 93943
thuynh@nps.edu

John S. Osmundson—John S. Osmundson received a BS in physics from Stanford University and a PhD in physics from the University of Maryland. He is an associate research professor with a joint appointment in the Systems Engineering and Information Sciences Departments at the Naval Postgraduate School in Monterey, CA. His research interest is applying systems engineering and computer modeling and simulation methodologies to the development of systems-of-systems architectures, performance models, and system trades of time-critical information systems. Prior to joining the Naval Postgraduate School in 1995, Dr. Osmundson worked for 23 years at Lockheed Missiles and Space Company (now Lockheed Martin Space Division) in Sunnyvale and Palo Alto, CA, as a systems engineer, systems engineering manager, and manager of advanced studies. Dr. Osmundson is a member of INCOSE.

John S. Osmundson
Naval Postgraduate School
Monterey, CA 93943
josmund@nps.edu



Abstract

Acquisition of a system-of-systems can be an all new acquisition of multiple systems that are intended to operate together as a system-of-systems. Much more common in the DoD is acquisition of one or more new systems that are intended to interoperate with existing systems as a system-of-systems with new capabilities. In either case, successful acquisition of systems-of-systems (SoS) necessarily depends on effective contracting structures and processes for systems-of-systems acquisition. In this paper, a set of system-of-systems issues that needs to be addressed in SoS acquisition is identified and the current findings in this on-going research are discussed. The findings suggest sustainment of extensive systems engineering efforts within the SoS acquisition and change to the existing contracting structures and process and organizational structures to maximize the probability of SoS acquisition success. The resulting changes will be applied to current and future DoD SoS acquisitions.

Introduction

No universal agreement on a definition of the term “system-of-systems” exists, but many definitions have common basic elements. Sage and Cuppan (2001) describe a system-of-systems (SoS) as having operational and managerial independence of the individual systems as well as of emergent behavior. Maier and Rechtin (2002) describe systems-of-systems as systems with emergent behavior that are operationally independent, managerially independent, evolutionarily developed, and geographically distributed. Boardman and Sauser (2006) describe one of the differentiating characteristics of an SoS as autonomy exercised by the constituent systems in order to fulfill the purpose of the SoS. Other definitions include operational and managerial independence and geographical separations of the constituent systems. Two characteristics of the types of systems-of-systems normally considered in the US Department of Defense (DoD) acquisition are that the constituent systems of an SoS are not chosen, but rather mandated to belong to the SoS and that the SoSs are usually bounded. An SoS can consist of to-be-developed systems, existing systems, or some combinations of new and existing systems.

SoS acquisition in the US DoD is faced with many challenges. Some SoS programs have faced technical and management challenges, if not failures. The US Army’s Future Combat System program (US Army, 2002) has a serious budget overrun (GAO, 2002; 2007). The US Coast Guard’s Integrated Deepwater System suffers from the lack of collaboration between contractors and the system integrators’ inability to impose decisions on them (GAO, 2006).

With an aim to develop approaches that can prevent such SoS acquisition programs from failing, Ghose and DeLaurentis (2008) look into “types of acquisition management, policy insights, and approaches that can increase the success of an acquisition in the SoS setting.” They investigate the impact of SoS attributes, such as “requirement interdependency, project risk, and span-of-control of SoS managers and engineers—on the completion time of SoS projects.” Ghose and DeLaurentis (2008; 2009) cite the following:

the common causes of failure (Rouse 2007) within SoS acquisition processes as: a) misalignment of objectives among the systems, b) limited span of control of the SoS engineer on the component systems of the SoS, c) evolution of the SoS, d) inflexibility of the component system designs, e) emergent behavior revealing hidden dependencies within systems, f) perceived complexity of systems and g) the challenges in system representation.



In their work, they analyze the effect of requirement dependency, span-of-control and risk profiles on, as a success metric, the total time to complete the project. For example, they find that acquisition process completes in 19 time-steps with low span-of-control, as compared to 12 time-steps with high span-of-control. The concept of span-of-control of engineers and managers is also addressed in the work in this paper, as it is related to both the pre-acquisition and acquisition phases of SoS acquisition.

Osmundson et al. (2007) address SoS acquisition issues and their resolution by modeling simulation, but with a focus on SoS systems engineering. These issues include initial agreement to operate as an SoS, SoS control, organization of the SoS, identifying SoS measures of effectiveness (MOEs) and measuring effectiveness, staffing, team building and training for SoS operation, identifying data requirements, identifying and managing interfaces, risk management, SoS testing and managing emergent behavior. Each of these issues is briefly discussed here. A detailed elaboration of these issues and their resolution by modeling and simulation are in (Osmundson et al., 2008).

The work captured in this paper attempts to answer this question: Can new contracting concepts be developed to aid in maximizing the probability of SoS acquisition success? The usual systems acquisition success criteria apply: performance, schedule, and budget—systems to be developed within a desired schedule and within a budget and to perform according to requirements. Briefly, contracting refers to the federal government and DoD contract management policy and guidance, roles and responsibilities in DoD contract management. A detailed elaboration of these contracting elements can be found in (Rendon & Snider, 2008).

This paper treats a realistic scenario of an SoS acquisition program represented in Figures 1 and 2. It is realistic in the sense that it reflects some current DoD SoS acquisition programs. Figure 1 shows three separate, autonomous, individual systems (System A, System B, and System C). These systems are currently being acquired (researched, developed, tested, produced, and deployed). Each system is managed by a government program office and a contractor performing in accordance with the requirements of an acquisition contract. In this scenario, during the course of the acquisition of each individual system, a new mission arises and requires an SoS that consists of the three systems to be built; the government thus adds a requirement that each individual system become part of the SoS acquisition program. Figure 2 reflects the new SoS acquisition program. In this paper, the discussion of the contracting structures and processes for SoS acquisition pertains to this scenario.

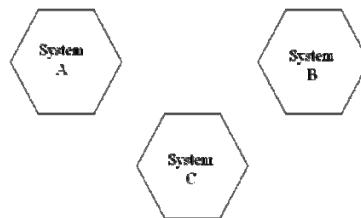


Figure 1. Three Separate Systems Being Developed

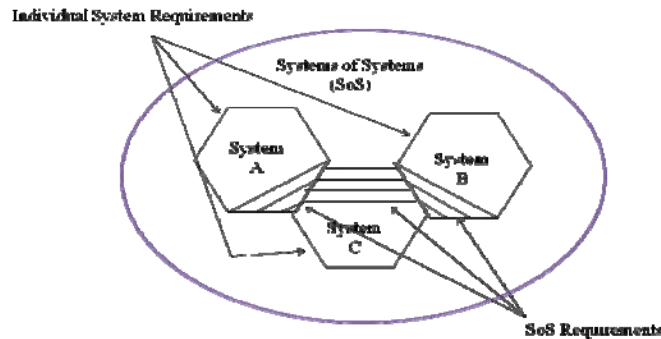


Figure 2. Addition of SoS Requirements

The transition from the acquisition of individual systems to the acquisition of an SoS has implications on the relationship between the government and the contractors. This relationship is also determined by the organizational structure used to manage the SoS acquisition program. Will the required SoS systems engineering be performed by a new, overarching group, by a collaboration among systems engineering organizations associated with existing systems, or by a single systems engineering organization associated with one of the component SoS systems? In addition to contracting, organizational structure is also discussed in this paper.

The goals of this paper are as follows:

1. To emphasize the span-of-control of engineers on SoS acquisition during the SoS pre-acquisition and acquisition phases,
5. To examine all possible contracting options in conjunction with all possible organizing options,
6. To arrive at the possible combinations of contracting and organizing options for resolving the SoS acquisition issues, and
7. To map resolution of SoS issues to the SoS acquisition success criteria.

The rest of the paper begins with a discussion of the SoS acquisition issues, follows with an examination of some SoS-acquisition-related concepts, and ends with a conclusion.

Recent System-of-Systems Acquisitions

Examples of recent SoS acquisitions are the US Army's Future Combat System, the US Coast Guard's Deepwater System, the Joint Tactical Radio System (JTRS) and Homeland Security's SBLnet, each of which has experienced technical, budget, and schedule challenges beyond what is considered the usual norm for single system acquisitions.

Future Combat System. The Future Combat System (FCS), shown in Figure 3, was originally to be composed of a networked system of new manned ground vehicles (shown on the right-hand side of Figure 3) and unmanned aerial vehicles (shown on the left side of Figure 3). FCS has recently been scaled back to a networked system of unmanned air and ground vehicles and existing manned ground vehicles.

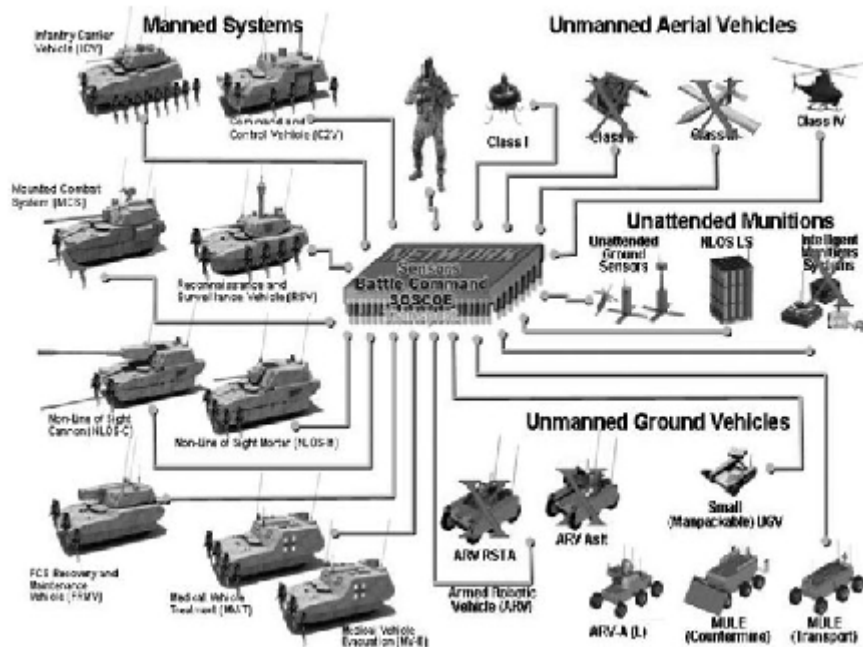


Figure 3. Original US Army Future Combat System Architecture

The initial program cost estimate was \$91.4 billion and the first combat brigade equipped with FCS was expected to roll out around 2015, followed by full production to equip up to 15 brigades by 2030 (Feickert & Lucas, 2009).

Deepwater. Deepwater, shown in Figure 4, originally was to include updated legacy ships plus new national security cutters, offshore patrol cutters and fast response cutters; updated aircraft and new manned and unmanned aircraft; and a new C4ISR system that would provide seamless communications between all of the assets, as show in Figure 5. The Deepwater program was begun in 2002, estimated to cost between \$19-24 billion, and expected to take 20-25 years to complete. The contract was awarded to Integrated Coast Guard Systems (ICGS), a joint venture of Northrup Grumman and Lockheed Martin, and ICGS hired subcontractors to design and build new assets. The Deepwater program was not only replacing old ships and aircraft, but was offering an integrated approach to upgrading other existing assets with improved C4ISR equipment and innovative logistics support systems (O'Rourke, 2009).



Figure 4. Original Coast Guard Deepwater Vessels and Aerial Vehicles



Figure 5. Networked Deepwater System-of-systems

The Deepwater program consisted of updating legacy assets and building new classes of cutters, such as the National Security Cutter, the Offshore Patrol Cutter, and the Fast Response Cutter; modernizing aircraft and building a comprehensive, long-term aviation force, including maritime patrol aircraft, unmanned aerial vehicles, and high-altitude endurance unmanned aerial vehicles; developing an integrated logistics support system; and modernizing the Coast Guard's command, control, communications, computers, intelligence, surveillance, and reconnaissance (C4ISR) systems to promote seamless communications between assets. C4ISR is fundamental to improving maritime domain awareness and is designed to not only ensure seamless interoperability among all Coast Guard units but also with Department of Homeland Security (DHS) components as well as with other federal agencies, especially the Navy.

Joint Tactical Radio System. The Joint Tactical Radio System (JTRS) is a software defined radio (SDR) that allows a single hardware platform to be reconfigurable so that it can accommodate multiple radio waveforms. JTRS accommodates legacy and new mobile ad hoc networking waveforms and can store and run multiple waveforms (Nathans & Stephens, 2007). JTRS is considered an SoS and consists of airborne-maritime fixed site (AMF) radios, ground mobile radios (GMR), handheld man pad small form fit radio (HMS), network centric enterprise services (NCES), GIG bandwidth extension (GIG-BE), and legacy networks. A model of the AMF delivery process is shown in Figure 6. Lockheed Martin was selected to serve as the Prime Systems Contractor (PSC).

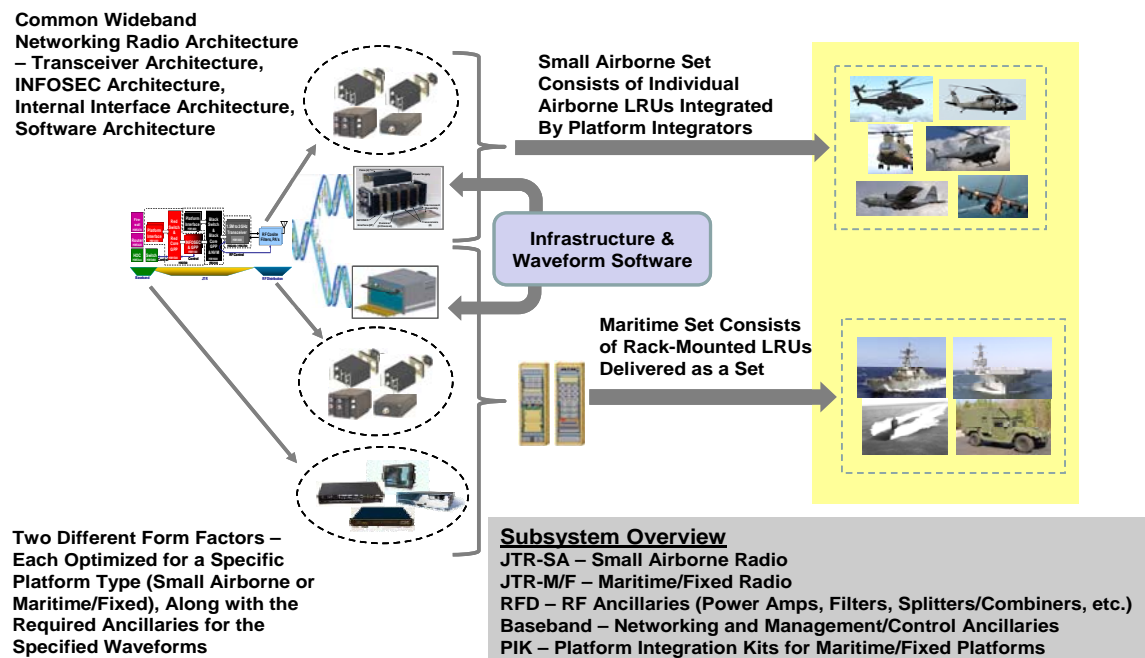


Figure 6. JTRS Airborne-maritime Fixed (AMF) Delivery Model
(JTRS, 2009)

SBlnet. SBlnet is a virtual fence designed to detect illegal crossings of the US southern border with Mexico. The virtual fence consists of a network of cameras, radars, lighting and other sensors—some mounted on elevated towers, as shown in Figure 7—and networked through a communication system that includes satellite nodes and links. The original contract was awarded to Boeing Integrated Defense Systems in 2006 and it was intended that the virtual fence would be in place, covering the entire US-Mexican border by 2011. At the time Boeing was awarded the contract, the cost was estimated to be \$2.5 billion (Montalbano, 2010).



Figure 7. SBInet Sensor Tower

Each of these SoS programs has followed a similar acquisition approach: contract to a large system integrator (LSI) who is supposed to have the responsibility and authority to manage the overall effort, including those of LSI subcontractors and vendors who number in a range from four to 100 or more. In the Lead System Integrator model, all four main tasks of building weapons have passed from government to industry. The LSI sets functional requirements and system specifications, provides program technical direction, controls program management, and controls program technical execution. The key reasons behind this shift are the growing complexity and scope of such systems while organic acquisition, systems engineering, program management resources within government are getting scarcer.

All of the example SoSs have experienced major challenges: two programs have been restructured with the customer organization taking over the LSI role from the private sector contractor; one program is in hiatus awaiting further investigation; and the fourth program faces almost certain large cost and schedule overruns.

FCS. There have been significant adjustments to the FCS program since its development start in 2003. The program was restructured and 4 of 18 core systems were cancelled. After the first four years of development, the Army estimated a total acquisition cost growth from \$91.4 billion to \$160.9 billion, while independent estimates were considerably higher—\$203.3 billion and \$233.9 billion. The program started with immature technologies and only 2 of the program's 44 technologies were fully matured by late 2006, according to the GAO, and it warned that all critical technologies may not be fully mature until the Army's production decision in February 2013. Requirements for networks and software were late, poorly defined or omitted due to the accelerated schedule for FCS development (Francis, 2008).

Deepwater. As an earlier part of the Deepwater program, the Coast Guard initiated an effort to modernize its existing 110-foot Island class patrol boats so that they could remain in service pending the delivery of replacement Deepwater craft. Among other things, the modernization increased the length of the boats to 123 feet. The effort is thus referred to variously as the 110-foot modernization program, the 123-foot modernization program, or the 110/123-foot modernization program. The initial eight boats in the program began to develop significant structural problems soon after completing their modernizations. The

Coast Guard removed the boats from service and canceled the program, having spent close to \$100 million on it. There were serious problems in the C4ISR system.

The Coast Guard is now pursuing Deepwater acquisition programs as individual programs, rather than as elements of a single, integrated program. The Coast Guard states that it is still using a systems approach to optimizing its acquisition programs, including the Deepwater acquisition programs, but that the system being optimized is now the Coast Guard as a whole, as opposed to the Deepwater subset of programs.

The Coast Guard announced in April 2007 that it would assume the lead role as systems integrator for all Coast Guard Deepwater assets. The Coast Guard is phasing out its reliance on ICGS as a LSI for Deepwater acquisition and will terminate the contract with ICGS in January 2011. To support its shift to that of the systems integrator, the Coast Guard is increasing its in-house system-integration capabilities.

JTRS. Since its initiation in 1997 until restructuring in 2006, the JTRS program experienced cost and schedule overruns and performance shortfalls, due primarily to immature technologies, unstable requirements, and aggressive schedules (GAO, 2006). More recently, the JTRS held a stakeholders review in December 2009, after several postponements of a scheduled CDR. Some of the recently identified issues are as follows: the current baseline relies on airborne platform processors to perform many of management functions, and while the platform processor will perform rudimentary radio control functions necessary to meeting the platform mission, relying on the platform processor for performing network management functions is unacceptable; JTRS is having some difficulty meeting NSA information assurance requirements; there have been a large number of requirements allocated by the LSI from upper levels to lower levels and not accepted by subcontractors at the lower levels; there is concern that some waveforms are not ready to be ported to JTRS; the current Platform Integration Kit (PIK) design doesn't integrate onto some platforms and some platforms do not want to use a PIK at all; and the software design and architecture is not fully defined and the definition would need to include operationally relevant system threads that demonstrate end-to-end capability. The JTRS program has extended its schedule and will also likely cost significantly more than current budget estimates.

SBI net. A GAO report on SBI net released in March 2010 identified a flawed testing process, performance issues, and poor management as serious ongoing issues affecting the program. The Department of Homeland Security cut off funding for the program, pending further review. Test plans were poorly defined and plagued by "numerous and extensive last-minute changes to test procedures," according to the GAO report, and even when the system was tested, it performed poorly. Further, those overseeing the project failed to prioritize solving problems with the system and failed to conduct further tests. The report concluded that if the development and testing of the system were to continue in the same fashion, SBI net would not perform as expected and would take longer and cost more than necessary to implement.

The DHS had expected the entire SBI net project to cost \$6.7 billion, a readjustment from its original projected budget of \$8 billion. To date, the DHS has spent about \$720 million on current SBI net deployments since the project began in 2005. The project was originally scheduled for completion by 2014, but the technical glitches and delays outlined in the GAO report held up the project so that only a prototype of the final solution is currently in use on just one part of the border. Funding for the SBI net has recently been suspended, pending Homeland Security decisions.



Systems-of-Systems Acquisition Issues

Systems acquisition refers to the disciplined management approach for the acquisition of an individual system, such as a weapon system (aircraft, ship, missile, etc.) or an information technology system. The acquisition process involves the various activities related to the development, design, integration, testing, production, deployment, operations and support, and disposal of the system. Within the federal government, specifically the DoD, systems acquisition uses a program management approach to the management of these activities. This approach involves the use of a project lifecycle, which includes phases, gates, and decision-points, a project manager, and a project team (Rendon & Snider, 2008). This approach is envisioned to apply to SoS acquisition, but making use of some new concepts, discussed in this paper, as there are significant differences between systems and systems-of-systems (SoS) and these differences affect the nature of government contracting for the development of systems-of-systems. Such application requires understanding of the issues associated with SoS acquisition.

The aforementioned SoS acquisition issues raised in Osmundson et al. (2008) are now briefly discussed here. In this paper, the importance of SE endeavor ties to the SoS pre-acquisition and acquisition phases and to the contracting process is emphasized; that is, the span of control of the engineers is crucial in these SoS pre-acquisition and acquisition phases.

- Initial agreement refers to decision-makers initially getting agreement that an SoS meets some desirable objective. It is an issue in particular when the SoS involves systems from different organizations or services because establishing an initial agreement is contingent on quantifying the benefits and risks of the new SoS.
- SoS control must be established: Who will control the SoS and how it will be controlled. Each partner may lose some measure of control over its own systems in order to enable overall SoS control.
- Organizing is a key issue of how to organize for the development and operation of an SoS. An example is the systems engineering process: How are processes that interface with SoS processes established and monitored?
- Staffing, team building, and training refer to how an SoS will be staffed and operated. SoS operations must be planned for, the skills required for SoS operations identified, and personnel with the proper skills acquired and trained in SoS operations.
- Data requirements is an issue concerning sharing of classified and/or proprietary design information among the SoS partners, who must recognize and weigh a possible loss of their systems' operational superiority based on the shared classified or proprietary design information against the SoS benefits.
- Interfaces must be identified and managed. Common language, grammar and usage must be established (for information SoSs), configuration management invoked to assure common agreements are followed, and required information security levels identified and provisions made to assure meeting of security requirements.
- Risk management at the SoS level is an issue related to the mitigation of SoS risks potentially effected by component systems, which requires detailed knowledge of component system risks and variations in individual system outputs.



- SoS testing requires that each SoS partner's system be tested in a manner that resolves any of its concerns about operational behavior and that SoS threads be tested.
- Measures of effectiveness is an issue because their strong dependence on individual component systems' measures of performance requires an understanding of the latter, and this issue is related to the issues of data requirements and interfaces.
- Emergent behavior, exhibited by the SoS resulting from unknown interactions among the constituent systems or from its interaction with the environment, need to be collectively understood, analyzed, and resolved, in particular when an emergent behavior may be detrimental to one or more of the partners.

Some SoS-acquisition-related Concepts

What contracting and organizing options can be used to aid in resolving the SoS issues? This section discusses these options and the correspondence of their combinations and the SoS issues.

Cross-functional Team Model. As previously stated, government systems acquisition management involves the use of project teams. The project team is a cross-functional team, consisting of technical specialists from the various functional areas involved in the acquisition process. These functional areas typically include systems engineering, contract management, financial management, logistics, and others. The cross-functional team is led by the government program manager. The program manager has overall responsibility for the success of the acquisition project. Although the program manager has overall responsibility, the program manager may not have all of the authority needed to manage the program. For example, the contracting officer may have the specific authority to award and make changes to the contract. Most systems acquisition programs involve effort performed by a contractor with the contract managed by the government program office. The contractor will generally have its own program manager and cross-functional team managing the contract for the contractor. Daily communication and coordination between government and contractor program managers, system engineers, and contract managers is the norm in defense acquisition management (Rendon & Snider, 2008). This paper is focused on systems engineering and contract management of the cross-functional team.

SoS Systems Engineering. In order to support knowledge-based acquisition, there is a need for effective global SoS systems engineering before the start of the acquisition process. Prior to milestone A, and prior to the Material Solution Analysis phase that cumulates in milestone A, an assessment must be made of technology opportunities and resources as well as of user needs. Assessment of technology opportunities and resources requires a global understanding of the proposed SoS and its operational environment. A technology may be considered mature when used in an existing system, but may lack required maturity when the existing system is incorporated into the proposed SoS and must operate under new conditions. An example is an information systems technology that is mature and stable when operating within the boundaries of a single system, but lacks the ability to interoperate with other systems.

Technology maturity assessment can also be considered one aspect of risk assessment and in the same way that technology maturity assessment must be made in the global context of the SoS, so must risk assessments.



The SoS must be represented in the pre-milestone A phase clearly, and in enough detail, to elucidate SoS technology and risk issues. A clear and complete SoS representation also elucidates data requirements and data ownership issues that will impact contractual relationships. SoS representation is a system architecting task that drives other early SoS systems engineering analyses and requires a very high level of skill. P.C. Lui, INCOSE Fellow and retired Singapore Defence Chief Scientist, remarked that while there are a limited number of good systems engineers, there is a very small number of systems architects. Yet, experienced, successful government and industry systems architects are essential at the start of SoS acquisitions. Good systems architecting will not assure program success, but lack of good systems architecting will almost always result in program failure.

One systems architecting approach is to represent SoSs in an object-oriented manner, using Systems Modeling Language (SysML), for example Huynh and Osmundson (2007) and Osmundson et al. (2004). Since SoSs can be represented in an object oriented modeling language, testing SoSs can be considered to be similar to integration testing of object-oriented software systems (Binder, 2000). System A is tested first and then a System B that interacts with System A is integrated with A and the combination is tested; then, a System C that interacts with A is integrated with A and tested; then, if Systems B and C interact together, B and C are integrated and tested, and then A, B and C are integrated and tested. These integration tests are based on thread of operations analysis, a part of the front end systems architecting process.

Knowledge of the availability of all systems is required early in the acquisition process in order to develop accurate test plans and program schedules. If a system is unavailable for testing, then a stub or driver is required; stub and driver development require complete knowledge about the missing system.

Thus, prior to milestone A and during the technology opportunities and resources assessment, there must be an SoS systems engineering team in place that has the high-level skills necessary to develop accurate SoS architectural representations, conduct technology assessments and risk assessments. High-level systems engineering expertise and systems engineering activities are necessary in order to assure knowledge-based acquisition. Without them, the SoS acquisition would begin with incomplete and possibly inaccurate technology maturity knowledge and risk knowledge.

The concept of span of control on the system components is crucial in all phases of acquisition. This means that systems engineering discipline need be enhanced and ever present in the SoS pre-acquisition and acquisition phases. Toward this end, there are two possible approaches. One is having a capable SE organization strictly organic to the SoS acquisition program office, and the other is using a capable SE organization external to the SoS acquisition program office, but the latter has strict ownership of the SE organization during the entire SoS acquisition. The advantages of the first approach are that the span of control of the engineers takes hold, direct control or exchanges are facilitated, and independence from contractors' undue influence materializes. The disadvantages are investment in money and people. The second approach suffers from control and increase in budgets for the same service required of the former, and time spent on establishing contracts to have an external organization to support.

Whereas this concept is not new, this paper calls for it to be instituted and for the span of control to exist during the pre-acquisition and acquisition phases.



Contracting Options. The transition from the acquisition of individual systems to the acquisition of an SoS has implications on the relationship between the government and the contractors. This relationship is largely determined by the contracting structure and processes governing the SoS requirements. There are three options for incorporating the SoS requirements into the individual acquisition programs (Programs A, B, and C in the scenario): two separate contracts, replacement of the existing contract, and modification of the existing contract. The discussion of each of them follows.

The first option is to incorporate the SoS requirements (shaded areas of each system in Figure 2) as a contract distinct from the existing contract for each contractor. Contractors A, B, and C would receive an additional contract with the specific SoS requirements for that specific system. In this option, each contractor would be working under two different and separate contracts—one for the acquisition of the basic system, and one for the SoS requirements related to the basic system.

The second option is to terminate the original contract for the acquisition of the individual system and to negotiate and award a new single contract for both the acquisition of the single system and the acquisition of the SoS components of that system. In this option, each contractor remains with only one contract.

The third option is to negotiate a modification to the existing contract, which incorporates the SoS requirements for that system under the existing contract. In this option, the contractor also remains with a single contract, albeit a modified contract, for all acquisition requirements.

This paper suggests that the third contracting option, modifying the existing contract to incorporate the SoS requirements, would be preferred over the first option, since having a contractor work under two separate contracts may be problematic. For example, there is a risk that the two contracts may be in conflict with each other, such as conflicting specifications, statements of work, or schedule priorities. The resources required for administering two separate contracts would be a disadvantage. Furthermore, managing two separate contracts would complicate organizational structures (discussed below). The third option would be preferred over the second option because modifying an existing contract is more advantageous than negotiating a termination agreement on the original contract and then negotiating a new contract with the contractor. During these negotiations, it is likely that the contractor would need to stop the acquisition effort, thus impacting the project schedule and cost.

Organizational Structure Options. Different SoS acquisition contracting options bear some impact on SoS acquisition program organizational structures. As previously stated, the transition from the acquisition of individual systems to the acquisition of an SoS has implications on the relationship between the government and the contractors. This relationship is also determined by the organizational structure used to manage the SoS acquisition program.

In structuring the organization, three options can be used for the SoS acquisition program. The first option is to designate one of the individual programs as the lead program and make that government program office responsible for managing the entire SoS acquisition program, which includes the other two systems. For example, the government program office managing System A could be designated the lead program and made responsible for ensuring systems (A, B, and C) meet the SoS requirements. Thus, the government program manager for System A will also have SoS acquisition responsibility and authority over the government program managers for System B and System C.



The second option is to establish a separate government program office responsible for the SoS acquisition program. This separate government program office would have SoS acquisition responsibility and authority over the three individual government program offices managing their individual acquisition programs (System A, System B, and System C.) In this option, the SoS acquisition management would be performed by in-house government acquisition and contracting workforce.

In the third option, a contractor is selected to manage the acquisition of the SoS program. This contractor, typically referred to as a Lead Systems Integrator, would oversee the SoS requirements within the three individual systems (A, B, and C). This option entails awarding a contract to a company to perform the SoS acquisition management.

This paper suggests that the second organizing option, establishing a separate government program office responsible for the SoS acquisition program, would be preferred over the first organizing option, since having one of the individual programs as the lead program and making that government program office responsible for managing the entire SoS acquisition program would result in potential conflicts of interest. The government program manager for the individual program may be biased and improperly influenced in the management of the overall SoS acquisition program. In this position, the government program manager may favor the individual program over the needs of the SoS.

The second organizing option would be preferred over the third organizing option because having a contractor manage the SoS acquisition program may involve the contractor performing some of the critical requirements determination and acquisition decision-making of the SoS program. The third contracting option may result in the outsourcing of inherently government functions related to the acquisition of the SoS program. It may also result in the government's loss of a systems engineering core competency and capability for managing SoS programs.

Integrating Acquisition Management Processes. In addition to the contracting options discussed above, another SoS issue relates to the integration of the SoS contract requirements among individual contracts. SoS acquisition programs involve a high level of uncertainty, and thus, a high level of risk. Since many of the individual systems have evolving requirements, and these requirements are required to interface with other individual systems in the SoS, the level of integration needed in the acquisition process of each individual system, as well as the SoS, is very high. Additionally, in SoS acquisition programs, the use of Lead Systems Integrators (LSIs) or prime systems contractors overseeing subcontractors performing the majority of the acquisition effort, also adds to the high need of integration within the acquisition process. In these SoS acquisition programs, one of the critical challenges is integrating the cost, schedule, and performance elements within the individual contracts (which now include the SoS requirements).

Many agencies respond to the increased uncertainty and risk of systems-of-systems acquisition programs by trying to increase the specificity of the contract elements such as performance requirements, contract type, incentive, delivery schedule, and other terms and conditions. Other agencies attempt to increase the flexibility of the contract elements to reflect the high uncertainty and risk. This flexibility would be reflected not in the detailed product or performance specifications of the contract, but more in the processes established for development of the specifications, testing and acceptance criteria, and cost allowability (Brown, Potoski & Van Slyke, 2008). A preferred approach is to strike a proper balance between contract element specificity and flexibility. This can be done through the development of an integrated management system, which will be discussed next.



In integrating the major elements of the SoS acquisition cost, schedule, and performance objectives, a best practice is to establish an integrated management system that integrates the planning, monitoring and control, and feedback elements of the SoS acquisition program.

The planning elements include the requirement specification, work breakdown structure, statement of work, and the integrated master plan. The integrated master plan (IMP) reflects all program activity, expands on the statement of work tasks, and defines the milestones. The IMP also specifies the program events, significant accomplishments, accomplishment criteria, and detailed tasks. The IMP is incorporated in the contract, along with the specifications, WBS, and SOW. However, it should be noted that the IMP is an event-driven plan and does not specify any calendar schedule. The JTRS AMF contract includes an IMP along with a Statement of Objectives, WBS, Performance Requirements Document, and other specifications.

The monitoring and control elements include the integrated master schedule (IMS), technical performance measures, and the earned value management system. Although not part of the contract, the integrated master schedule provides the detailed calendar schedule for tracking schedule progress; the earned value management systems tracks cost and schedule performance; and the technical performance measurement system tracks the technical risk. The JTRS AMF program includes an IMS as well as technical performance measures and an earned value management system.

The feedback elements include the contract award fee, if any, and contractor performance assessment reviews. The contract award fee allows the contractor to earn additional profit, based on over and above required levels of performance. The contractor's performance and any award fee decisions are based on a subjective evaluation of the government. The contractor performance assessment review is separate from the award fee and applies to all government contracts exceeding the simplified acquisition threshold. The JTRS AMF contract includes an award fee for the design, development, delivery, and testing of the Engineering Development Models.

This integrated management system would be developed and used for each individual system acquisition program, as well as developed and used at the SoS level by the government program office responsible for the SoS acquisition program, as discussed in the previous organizational structure options.

Linkages between Contracting Options and Organizational Structure

Options. A logical linkage appears to exist between the preferred contracting and organizing options for transitioning from the acquisition of individual systems to the acquisition of an SoS. The preferred contracting option of modifying the existing contracts to incorporate the SoS requirements and the preferred organizing option of establishing a separate government program office responsible for the SoS acquisition program can be effectively implemented together. The government program office responsible for the acquisition of the SoS would be the requirements agency for the SoS program. In this capacity, the SoS government program office could communicate the SoS requirements to each system program office. The system program office would then incorporate these requirements into the individual system contract modification. The systems engineering and contract management personnel from the SoS government program office would communicate and collaborate with the systems engineering and contract management personnel in each of the individual system program offices to manage these SoS requirements.



One potential drawback to the linkage of the two preferred contracting and organizing options would be the conflict potential between the SoS government program manager and the individual system government program manager (such as between the SoS government program manager and System A government program manager). This would occur in situations dealing with cost, schedule, and performance priorities between the two aspects of the system (individual and SoS). The understanding of and adherence to roles and responsibilities between the SoS government program manager and the individual system program manager, as well as an order of precedence clause in the contract, would help deter these potential conflict situations.

Table 1 shows a number of possible combinations of contracting and organizing options, which (marked with “√”) potentially result in the resolution of the SoS issues, which, in turn, enable satisfaction of the SoS acquisition success criteria (marked with “X”). As discussed above, the preferred contracting option for the scenario of interest is the replacement of the existing contract. It can be combined with either the separate government program option, which is, as discussed above, the preferred option or the lead systems integrator option. For example, given that the existing contract is replaced by a new one, either the separate government program option or the lead systems integrator option, the SoS interfaces issue should be resolved. The resolution of such an issue would enable the satisfaction of the SoS acquisition criteria.

Table 1. Resolution of SoS Issues by Option Combinations and Satisfaction of Acquisition Success Criteria

Issues	Contracting Option			Organizing Option			Acquisition Success Criteria		
	Two separate contracts	Replacing contract	Modified contract	Designated individual program	Separate government program	Lead Systems Integrator	Performance	Schedule	Budget
Initial agreement .		√			√	√	X		
SoS control					√	√	X		
Organizing		√			√	√	X	X	X
Staffing, team building, and training		√			√				X
Data requirements		√			√		X	X	
Interfaces		√			√	√	X	X	X
Risk management		√			√	√	X	X	X
SoS testing		√			√	√	X	X	X
Measures of effectiveness		√			√	√	X	X	X
Emergent behavior		√			√	√	X		

Conclusion

The purpose of this on-going research is to determine contracting and organizational options to enable successful SoS acquisition and to apply them to current and future DoD SoS acquisitions.

At this point in this research, the following is suggested:

- A sustainable systems engineering effort with an extensive span of control by systems engineers within an SoS acquisition is necessary for a successful SoS acquisition.

- Among the possible contracting options, replacing the contract is the preferred option. But that's not sufficient. Organizing options must be considered, for an organizing option must be coupled directly with a contracting option and, together, they would enable resolution of the SoS acquisition issues, which, in turn, could improve the probability of SoS acquisition success, thereby facilitating and effectively managing the SoS acquisition effort.

These findings will be applied to a case study, whose results will be published in a future paper. Furthermore, a separate paper will invoke a collaboration theory and incorporate it in the organizing options in particular and in SoS acquisition in general (Huynh et al., 2010).

References

- Binder, R. (2000). *Testing object-oriented systems: Models, patterns, and tools*. Addison-Wesley.
- Boardman, J., & Sauser, B. (2006, April 24-26). The meaning of system of systems. In *Proceedings of the IEEE International System of Systems Conference*. Los Angeles, CA.
- Brown, T.L., Potoski, M., & Van Slyke, D.M. (2008). *The challenge of contracting for large complex projects: A case study of the Coast Guard's Deepwater Program*. Washington, DC: IBM Center for the Business of Government.
- Chen, P., & Clothier, J. (2003). Advancing systems engineering for systems-of-systems challenges. *Systems Engineering*, 6(3), 170-183.
- Chuen, L.P. (2007, March). Chief Defence Scientist, Ministry of Defence, Singapore. [Personal communication].
- Feickert, A., & Lucas, N.J. (2009, November 30). *Army future combat system (FCS) "spin-outs" and ground combat vehicle (GCV): Background and issues for Congress*. Congressional Research Service Report to Congress (7-5700 RL32888). Washington, DC: US GPO.
- Francis, P.L. (2008, April 10). *2009 review of future combat system is critical to program's direction* (GAO-08-638T). Washington, DC: US GPO.
- GAO. (2002, July 1). *Issues facing the Army's future combat systems*. Retrieved July 1, 2002, from <http://www.gao.gov/new.items/d031010r.pdf>
- GAO. (2006, June). *Coast Guard: Observations on agency performance, operations and future challenges*. Retrieved from <http://www.gao.gov/new.items/d06448t.pdf>
- GAO. (2006, September). *Restructured JTRS program reduces risk, but significant challenges remain* (GAO-06-955). Report to Congressional Committees. Washington, DC: Author.
- GAO. (2007, January 2). *Future combat system risks underscore the importance of oversight*. Retrieved January 2, 2007, from <http://www.gao.gov/new.items/d07672t.pdf>
- Ghose, S., & DeLaurentis, D.A. (2008, May 14-15). Defense acquisition management of systems-of-systems. In *Proceedings of the Fifth Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School, Monterey.
- Ghose, S., & DeLaurentis, D.A. (2009, July). Development of an exploratory computer model for acquisition management of system-of-systems. In *Proceedings of the INCOSE Symposium*. Singapore.
- Huynh, T.V., & Osmundson, J.S. (2007, March 23-24). An integrated systems engineering methodology for analyzing systems of systems architectures. In *Proceedings of the 1st Asia-Pacific Systems Engineering Conference*. Singapore.



- Huynh, T.V., Osmundson, J.S., & Rendon, R.G. (2010, forthcoming). *On collaboration in systems-of-systems acquisition*.
- JTRS. (2009, December 2-3). *Stakeholder review*. San Diego, CA.
- Maier, M., & Rechtin, E.I. (2002). *The art of systems architecting*. Boca Raton, FL: CRC Press.
- Montalbano, E. (2010, March 19). GAO: Multiple failures sunk border security system. *Information Week*.
- Nathans, D., & Stephens, D.R. (2007, July). Reconfiguring to meet demands: Software-defined radio. *Crosstalk Magazine*.
- O'Rourke, R. (2009, December 23). *Coast Guard Deepwater Acquisition Programs: Background, oversight issues, and options for Congress*. Congressional Research Service Report for Congress (7-5700 RL33753). Washington, DC: US GPO.
- Osmundson, J.S., Gottfried, R., Kum, C.Y., Boon, L.H., Lian, L.W., Patrick, P.S.W., & Thye, T.C. (2004). Process modeling: A systems engineering tool for analyzing complex systems. *Systems Engineering*, 7(4).
- Osmundson, J.S., Langford, G.O., & Huynh, T.V. (2007, March 23-24). System of systems management issues. In *Proceedings of the Asia Pacific Systems Engineering Conference (APSEC)*. Singapore.
- Rendon, R.G., & Snider, K.F. (Eds.). (2008). *Management of defense acquisition projects*. AIAA, Library of Flight Series, Reston, VA.
- Sage, A.P., & Cuppan, C.D. (2001). On the systems engineering and management of systems-of systems and federations of systems, information, knowledge. *Systems Management*, 2(4), 325-345.
- US Army. (2002). *Future combat systems*. Retrieved from <https://www.fcs.army.mil/>



Creation of a System of Systems Portfolio Management and Technology Selection Methodology

Richard Volkert—Mr. Richard Volkert is employed by the Space and Naval Warfare (SPAWAR) Systems Center-Pacific as the ISR Departments Lead Systems Engineer supporting SPAWAR and as the PMS 420 Deputy Technical Director. Mr. Volkert has over 27 years of service in the government, including twenty years as an active duty Naval officer with service as an engineering duty officer and in submarines. Over 19 years of that time he has been involved in the fields of research, development, acquisition, and systems engineering. He possesses degrees in Aerospace Engineering and Acoustical Engineering and is presently enrolled in a PhD program for Systems Engineering.

Rich Volkert
SPAWAR Systems Center -Pacific
Commanding Officer, SSC-Pacific, 53560 Hull St (BS Bldg 1 Rm A526), San Diego, CA 91977
richard.volkert@navy.mil
Phone: 619 553-1616

Carly Jackson—Ms. Carly Jackson is a Systems Engineer at SPAWAR Systems Center Pacific. She is currently supporting the Littoral Combat Ship Mission Module Program Technical Director's office, primarily tasked with the implementation of product and process commonality across complex Systems of Systems and Systems Engineering preparations for the impending Milestone B. Ms. Jackson earned simultaneous BS and MS degrees in Mechanical Engineering from UCLA in 2002 and an MBA in Business Administration from Pepperdine University in 2007. She is Level I certified by DAU in SPRDE-Systems Engineering, Test and Evaluation, Program Management, and Production, Quality, and Manufacturing.

Carly Jackson
SPAWAR Systems Center -Pacific
Commanding Officer, SSC-Pacific, 53560 Hull St (BS Bldg 1 Rm B519), San Diego, CA 91977
carly.jackson@navy.mil
Phone: 619 553-1492

Peter Gentile—Mr. Peter Gentile is an Associate Systems Engineering Fellow for the Northrop Grumman Corporation, currently working as the Technical Director and Chief Engineer for the Littoral Combat Ship Mission Package Integration program. His 42 years of engineering includes wide involvement in the design of electronics, computers, software, and systems for platforms covering sea to space. Mr. Gentile earned a Bachelor's degree in Electronics Technology from NY Institute of Technology and a Master's in Systems Engineering from The Stevens Institute of Technology in Hoboken, NJ.

Lance Harper—Mr. Lance Harper is employed by the Northrop Grumman Corporation, where he works as a Systems Engineer supporting the Littoral Combat Ships Mission Package Integrator (LCS MPI) contract. Mr. Harper has been working for 10 years on government contracts, focusing on research, development, and systems engineering. He earned a Bachelor's degree in Electrical Engineering from the Johns Hopkins University.

Lance Harper
Northrop Grumman Corporation
300 M St SE, Suite 100, Washington DC 20003
lance.harper@ngc.com
Phone: 202 799-3132



Art Van Nostrand—Mr. Art Van Nostrand is currently a systems engineer at Northrop Grumman Corporation, working on the Littoral Combat Ship (LCS) Mission Package Integrator (MPI) Program. Mr. Van Nostrand has over 30 years of systems engineering experience. He has worked on both DoD and civilian systems development. He has a BS degree in Physics from Long Island University, an ME in Systems Engineering from Stevens Institute of Technology, and an MBA in Finance from Dowling College. Mr. Van Nostrand is a member of the International Council on Systems Engineering (INCOSE), and is a Certified Systems Engineering Professional (CSEP) with the DoD Acquisition Extension.

Art Van Nostrand
Dept/MS (03790/W28-035), 925 South Oyster Bay Road, Bethpage, N.Y. 11714-3582
art.vannostrand@ngc.com
Phone: 516 575-3794

Tom Sondi—Mr. Tom Sondi is an Engineering Manager for the Northrop Grumman Corporation, currently working as the Systems Engineering Integration and Test (SEIT) lead for the Littoral Combat Ship Mission Package Integration program. His 30 years of engineering and management includes a wide involvement in systems engineering in the areas of requirements development, requirements analysis, system design, integration and test of systems for military platforms covering sea to space. Mr. Sondi earned a Bachelor of Engineering degree in Electrical Engineering from Manhattan College.

Tom Sondi
Northrop Grumman Corporation
Dept/MS (03674/W28-035), 925 South Oyster Bay Road, Bethpage, N.Y. 11714-3582
tom.sondi@ngc.com
Phone: 516 575-1646

Kenneth Michaud—Mr. Kenneth Michaud assumed his current responsibilities as Vice President of Pioneering Evolution in July 2009. Prior to joining Pioneering Evolution, Mr. Michaud was the Acquisition Manager for Littoral Combat Ship Mission Modules Program Office (PMS 420) under the PEO LMW from 2004 through his departure from Government service in July 2009. He served as a Subject Matter Expert in Program Management and Acquisition Management within the Government for over 18 years. Mr. Michaud is a graduate of Defense Acquisition University's Executive Program Management course (PMT 401) and the recipient of the Meritorious Civilian Service Award from the Department of the Navy.

Kenneth Michaud
Pioneering Evolution, LLC
1316 Independence Avenue SE, Washington, DC 20003
kmichaud@pioneeringEvolution.com
Phone: (202) 546-0096 Ext. 5

Abstract

As OSD seeks to field new capabilities while working to reduce cost and risk, it becomes imperative that systems engineering tools evolve. Traditionally, cost/schedule monitoring, technology assessment, and performance analyses have been conducted as independent activities focused on systems. However, as systems become more complex and entwined into operating as components of System of Systems (SoS), the need for more insight during the design and development stages increases. This dictates the need for a



methodology for SoSs that allows for fully integrated analysis and trade-offs of the technical, cost, and schedule design spaces.

The US Navy (PMS 420, SSC Pacific), Northrop Grumman, and the Stevens Institute of Technology are collaborating to develop such a comprehensive financial and portfolio management methodology for SoSs. The concept leverages the System Readiness Level (SRL) as a measure of SoS development and coalesces a system performance monitoring approach that provides insight into both current and anticipated performance. Additionally, a methodology for understanding the impact of technical trades within a SoS is introduced. Together, these tools allow for a true trade-off analysis capability that can be used to examine the extent to which a set of technology options either meet budget constraints or maximize performance.

The Challenges of System of Systems Management

The Department of Defense (DoD) has seen a growth in the acquisition of SoSs over the last few decades. This trend is expected to continue as the DoD increases focus on capabilities without changing its system oriented acquisition organization. While providing significant opportunities for extending mission capabilities through the integration of existing and new capabilities into a synergistic SoS, there exists significant systems engineering challenges related to the integration and management of SoS. These engineering challenges are discussed in the *Systems Engineering Guide for Systems of Systems* (ODUSD(A&T)SSE, 2008).

One example of the challenge presently facing SoS Program Managers (PMs) is the understanding of the SoS technical maturity. Historically, the Technology Readiness Level (TRL) methodology has been a key gauge of the technical maturity for individual systems within the DoD for the better part of two decades. However, when TRL is applied to components within a SoS, the model of using individual technology maturity as a measure of readiness for SoS development quickly breaks down. TRLs simply do not account for integration maturity or the complexity of bringing together any number of independent technologies to function as a SoS.

Similar problems also become apparent with many other systems engineering and program management tools when applied in a SoS context, including Technical Performance Measures (TPMs) as used to track progress toward achieving Key Performance Parameters (KPPs) and Earned Value (EV) Management (to track cost/schedule). Existing tools simply do not provide sufficient insight into SoS development, contributing to a rash of complex development and acquisition projects that have gone astray. In a 2006 study (GAO, 2006, September 14) the Government Accountability Office (GAO) noted that a lack of insight into the technical maturity of complex systems during development has contributed to an environment of significant cost overruns, schedules slips leading to program delays, canceled acquisition efforts, and reduced system performance at fielding. In case after case, failure is most commonly not found at the technology development level, but rather at the point of a combination of two or more elements.

This paper provides insight to the methodologies and tools being developed and used by the Littoral Combat Ship (LCS) Mission Modules Program Office (PMS 420) to assist the PM and his staff in meeting the SoS development challenge.



The Mission Modules Program—An Acknowledged SoS Example

Acknowledged System of Systems Definition

The DoD *System of Systems Engineering Guide* defines an acknowledged SoS as a set or arrangement of systems that results when independent and useful systems are integrated into a larger system that delivers unique capabilities. An acknowledged SoS has recognized objectives, a designated manager, and resources. Within an acknowledged SoS, the constituent Mission Systems (MS) retain their independent ownership, objectives, funding, development and sustainment approaches. Changes in the MSs are based on collaboration between the SoS PM and the MS PM. This complicates the task of a SoS PM and system engineer who must navigate the evolving plans and development priorities of the SoS constituent systems, along with their asynchronous development schedules, to plan and orchestrate evolution of the SoS toward SoS objectives.

The LCS Mission Modules Program as an Acknowledged System of Systems

The LCS Mission Modules Program Office (PMS 420) was established by the Assistant Secretary of the Navy for Research, Development, and Acquisition (ASN RD&A) on October 1, 2003 within the Program Executive Office—Littoral and Mine Warfare (PEO LMW) for the development, acquisition, and sustainment of the modular Mission Packages (MPs). The initial focus of PMS 420 was to take existing independent capabilities in the fields of surface warfare (SUW), mine countermeasure (MCM), and anti-submarine warfare (ASW) and to integrate and modularize those capabilities to provide deployable and swappable warfighting capabilities for the LCS. Thus, the LCS MPs meet the definition of being a SoS as they are made up of individual MSs, including vehicle, communication, sensor, or weapon systems; support equipment, including support containers, or vehicle cradles; software; mission crew detachments; and aviation systems; these are then integrated into a larger system to deliver unique capability. As the charter of PMS 420 is to acquire, integrate, modularize, and sustain focused warfighting capabilities from existing program lines, PMS 420 primarily serves as a Ships Acquisition Program Manager (SHAPM) with a focus on acquiring the individual mission systems from Participating Acquisition Resource Managers (PARMS) who manage existing product lines and programs of records. This lack of direct management responsibility for the individual mission systems means that the SoSs comprising the MPs are an acknowledged SoS.

Mission Package Explained

The hierarchal MP concept, illustrated in Figure 1, is best described in three layers.

These layers are:

- Mission System (MS) = a single Vehicle, Communication, Sensor, or Weapon System
- Mission Module (MM) = a combination of mission systems + Support Equipment + Software that provide a unique mission capability
- Mission Package (MP) = the collection of MMs + Mission Crew Detachments + Support Aircraft that provides the required ability to conduct a focused warfighting mission as required by the LCS Capability Development Document (CDD).



A MP consists of MSs integrated into warfighting responsive MMs, mission crew detachment and mission configured aircraft with their composite aviation crew detachment. MMs combine MSs (vehicles, sensors, weapons) and support equipment that install into the Seaframe via standard interfaces.

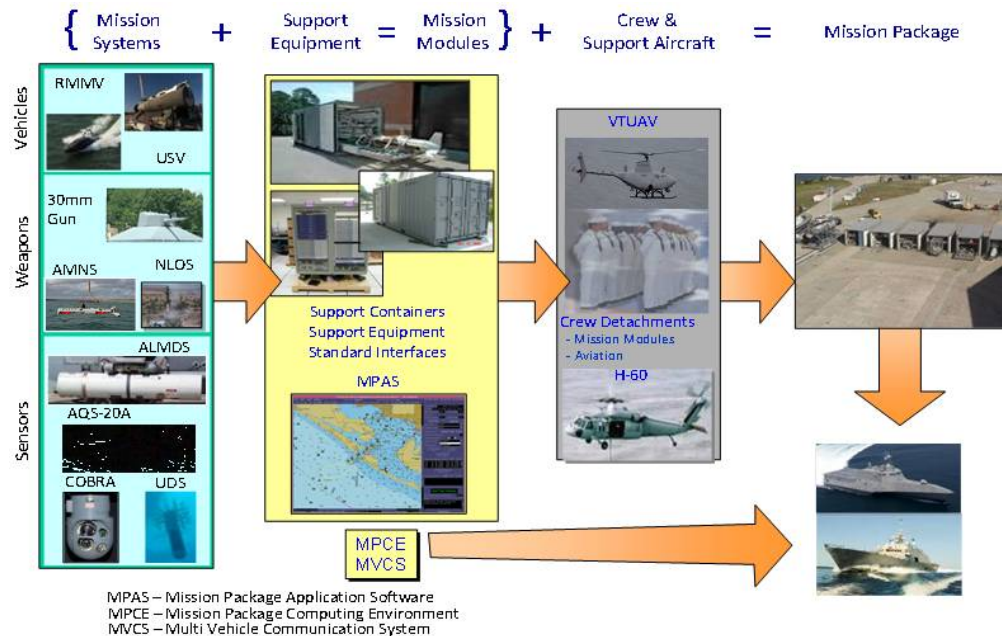


Figure 1. Mission Package Defined

MSs are sized to fit inside standard ten- or twenty-foot International Organization for Standardization (ISO) support containers (SCs), or on ISO compliant flat racks and vehicle cradles. Using ISO SCs simplifies shipping, storage, availability of correct handling equipment, and container movement from shore to ship and ship to shore as the Navy is able to leverage the intermodal transportation resources used for shipping commercial cargo worldwide. MP reconfiguration will occur in homeport or overseas, using pre-positioned MPs or MPs that have been transported into theater by air or sea and staged near the LCS operating area.

MPs can be swapped in order to reconfigure the ship for a different mission in a short period of time, giving a Combatant Commander a uniquely flexible response to changing warfighting requirements. To achieve this flexibility, the Navy is developing and procuring specific numbers of MPs to meet the Fleet's warfighting requirements. The quantity of each MP type differs based on analysis of projected operational needs.

Mission Package Status

The first two LCS ships, USS Freedom (LCS-1) and USS Independence (LCS-2), have been delivered to the Navy along with prototype MPs to the Navy for use in testing the designs and the concept of modular warfighting capability. The LCS MM Program (hereafter referred to as "the program") is presently proceeding to a Milestone B acquisition decision. The first three MPs (1 each of SUW, ASW, and MCM) have completed Design Readiness Review (DRR) and are engaged in further developmental testing of their capabilities either independently in End-to-End (E2E) testing and/or by being integrated onto the LCS and tested as part of the Seaframes integrated warfighting capability. The rollout for the first

MCM MP was on September 14, 2007. Phases 1 and 2 of E2E testing for MCM MP 1 were completed September 25, 2008, and August 19, 2009, respectively. Phase 3 of E2E testing for MCM MP 1 is scheduled for Summer 2010. The first ASW MP was rolled out on September 19, 2008. E2E testing for ASW MP 1 was completed April 3, 2009. ASW MP 1 Developmental Test (DT) is scheduled for Summer 2010. The first SUW MP was rolled out on July 11, 2008. E2E testing for SUW MP 1 was completed in July 2009. The first SUW MP was recently called on for an early deployment mission. In February 2010, LCS 1 deployed with the first SUW MP augmented with a prototype Maritime Security Module (MSM) to provide Visit, Board, Search, and Seizure (VBSS) capability to the Southern Command.

PMS 420 Progress on Addressing the Management Challenges for SoS Programs

PMS 420, since its founding, has recognized the challenge of leading an acknowledged SoS development and quickly began development of novel system engineering tools and methodologies, designed to ultimately reduce risk and provide enhanced management (technical, cost, schedule) insight into the SoS problem. Four initial areas of traditional programmatic concern have been developed to date. Lessons learned by PMS 420, approaches used and their benefits will briefly be discussed within this paper. These tools/processes include the areas outlined below.

1. **SRL Used to Determine SoS Maturity Analysis:** The System Maturity Model (SMM), developed by the Stevens Institute of Technology (SIT) and Northrop Grumman under funding provided by PMS 420. The SMM has been applied for the purpose of monitoring the maturity and integration status of individual technologies within the MP SoSs for PMS 420 and will be discussed in Section 3.1.
2. **Requirements Management & the Drive towards Commonality in a SoS World:** Requirements management is a significant challenge within an individual system development effort. This becomes even more challenging within the interrelated development environment of the three SoSs that PMS 420 faced. As an acquiring PM, PMS 420 has limited ability to impact Life Cycle Cost (LCC) of acquired capabilities. However, through implementation of an effective centralized and consolidated requirements management capability, PMS 420 is using commonality as a means to drive down cost. This approach will be expanded upon in Section 3.2.
3. **Expanding Financial Management past EV:** Financial and task management within a SoS is a complex task. PMS 420 required new processes and tools that could improve the PM's ability to monitor and review task execution and earned value, support multi-year pre-planning of research, development, procurement, and sustainment efforts at warfare centers and contractors. In addition, there was a need for greater insight into the cost of risk management activities and a desire to reduce funding document touch-time and rejections rates. PMS 420 has developed such a tool, and a description of the tool and approach used by PMS 420 to accomplish this will be expanded upon in Section 3.3
4. **Understanding and Influencing SoS Reliability:** Traditionally, system reliability has been determined through the calculation of independent critical component reliability in the system and defined by the value of Operational Availability (A_o). In a SoS, especially in a mission-focused area, this approach



may no longer be the best metric for use. PMS 420 has been evaluating an approach based on mission completion and the use of reliability block diagrams developed to represent mission strings, which will be expanded upon in Section 3.4.

In addition to the tools that have been developed to date, two additional areas will be presented that PMS 420 is presently investigating that are designed to further enhance the PM's ability to gain insight into understanding the status and risks associated with SoS development. These tools, while still under development and evaluation by PMS 420, are designed to assist the SoS PM in the areas of conducting capability tradeoffs and in the understanding of the ability to achieve required performance. Specifically, these tools are designed to provide insight into areas listed below.

1. **Evaluating the Impact of Technology Insertion:** One of the strengths envisioned from the MP SoS is their perceived ability to adapt to and rapidly and effectively incorporate new technologies that can provide increased warfighting capabilities. If the integration and maturation risks are not fully understood, a perceived improvement could actually lead to a decreased capability or significantly increased programmatic costs. To avoid these negative effects, PMS 420 has been developing a methodology to assess the impact of technology insertions in support of conducting tradeoff analysis. This proposed approach will be expanded upon in section 4.1.
2. **Predicting SoS Performance:** One of the challenges of an acknowledged SoS, is that the SoS PM may have little ability to monitor performance development of the individual systems. This issue is further complicated by the fact that individual system performance, when integrated into a SoS, may be different. How, then, can the SoS PM determine if they are on track to satisfy the KPPs for the SoS? PMS 420 has been developing a Performance Level Monitoring Model that seeks to provide the PM with this form of insight. The proposed approach will be expanded upon in section 4.3.

Systems Engineering Management and Insight in SoS Programs

As discussed in the previous sections of this paper, a SoS usually does not directly control the development of the majority of the technologies comprising the acknowledged SoS. This is a common situation and poses significant challenges to management in understanding where the end capability of a System stands with respect to providing the required level of performance as specified in the KPPs and the TPMs and in understanding the level of developmental and integration risk of the individual technologies composing the SoS. To resolve this area of engineering management concern, PMS 420 started the development of a portfolio of SoS Management tools. This section of the paper discusses some of these tools and lays the foundation for the later discussion of future efforts.

SRL Used to Determine SoS Maturity Analysis

LCS MPs will deliver required capability via the fielding of a series of incremental MPs until full capability that satisfies the LCS CDD is reached. For example, Increment 1 of the MCM MP provides capability for the detection and neutralization of volume and bottom mines. Increment 2 of the MCM MP introduces inshore detection capability via the Coastal Battlefield Reconnaissance and Analysis (COBRA) system. Increment 3 introduces



additional MCM capability to the Fleet, including a magnetic and acoustic sweep capability to address the bottom/buried mines threat using the AN/ALQ-220 Organic Airborne and Surface Influence Sweep (OASIS) system. The full MCM baseline capability will be achieved by Increment 4 in FY17, which introduces the AN/AWS-2 Rapid Airborne Mine Clearance System (RAMICS), that neutralizes near surface and floating mines.

As presented to this symposium last year, in order to gain insight and manage the development maturity of these incremental deliveries, PMS 420 has implemented an emerging concept known as the SRL (Forbes, Volkert, Gentile & Michaud, 2009). By pairing the traditional TRL scale with a new series of criteria known as the Integration Readiness Level (IRL), a more complete look at true system maturity can be obtained (Sausser, Ramirez-Marquez, Magnaye & Tan, 2008). Under this methodology, the readiness of each technology is still considered, but instead of being a stand-alone metric for determining readiness for incorporation, it is analyzed in concert with both its integration requirements and the maturity of other technologies with which it interfaces. The calculation of SRL is described in the above-referenced papers. The SRL methodology has been highly successful on the program and has paid dividends in terms of both increasing decision-maker visibility into true system status and allowing for pre-emptive actions to be taken to mitigate potential developmental issues.

An example of the use of this approach by PMS 420 in understanding the impact to the program of technology options is shown in Figure 2, where the SRL for the initial configuration (on the left side of the figure) of the MCM MP number 1 is calculated as 0.57. In this configuration, one system component, the Multi-Vehicle Communications System for the Remote Multi-Mission Vehicle (MVCS (RMMV)), is early in its maturity and is lagging most other components, both in its technology and its integration readiness. This configuration resulted in a lower than acceptable overall SRL of 0.57, beyond the risk threshold of the program.

The program evaluated the replacement of this lagging component with the combination of a Data Link System (DLS) both on-board and on the RMMV, each of which have both better TRLs and IRLs. This is shown on the right side of Figure 2. In this manner, the overall SRL of the MCM MP 1 increased to 0.67, now within the range acceptable to the established risk threshold of the program.

The program has used this methodology to monitor developmental status by incorporating it into a continuing quarterly evaluation of the SRL level for each of the mission packages. This consistent evaluation allows the PMS 420 PM to better understand maturation of the individual MP SoS and of each increment within the SoS. In turn, this provides him with a greater understanding of the program's technical status, enabling the PM to better maintain and manage the development risk of the MPs as they progress through design and development.



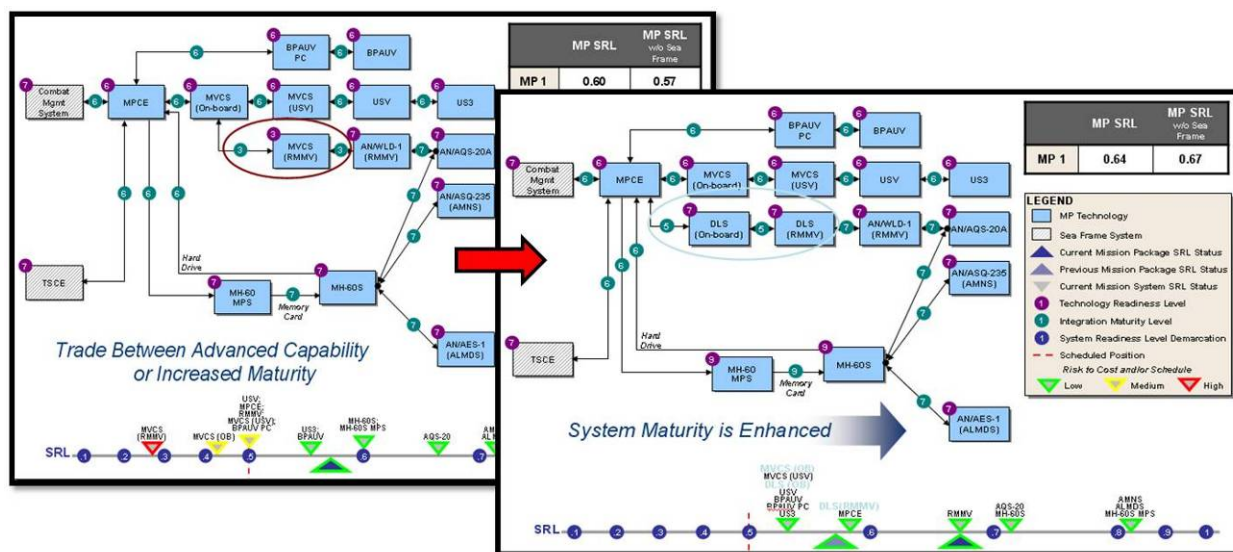


Figure 2. Initial and Enhanced Readiness of MCM MP 1

An example of the progress of the SRL level for two increments of the SUW MPs is illustrated in Figure 3 for the SUW MP. The graph demonstrates how the SRL value of different increments can be affected by both similar and [different] systems that make up the increments. The main dip in the graph (4-5) indicates a problem that was identified with a shared system between the increments discovered during interface testing, while the rise in the graph (5-6) demonstrates the results of correcting that identified problem. The lower overall SRL of the SUW Increment 2 is associated with the lower maturity level of one its component technologies, which is being consistently monitored by PMS 420.

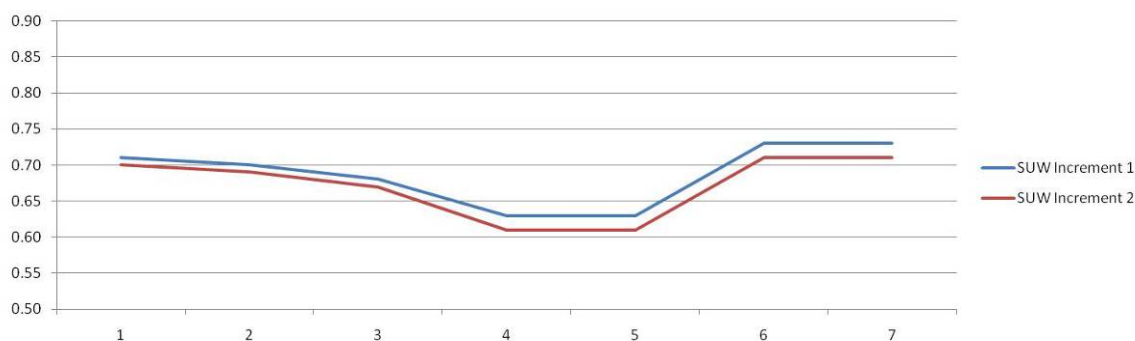


Figure 3. SRL Values Over Time

Since the initial presentation of the SRL method, the program has developed and documented a comprehensive process for System Maturity Assessments (SMA) and described its application to generic SoSs. The SMA process is iterative with a structured set of well-developed tasks that are described in detail in the *System Maturity Assessment Guide* (PMS 420, 2009) and illustrated in Figure 4. The first three steps of this process need only be conducted during initial system architecture development. Once the system architecture and subsequent system designs have been placed under configuration control, successive assessment iterations need only review the previous TRL and IRL criteria for any updates due to development progress and then recalculate the SRL with updates to reporting mechanisms conducted as needed. The fundamental basis of the SMA process is

the proper creation of an assessment framework to include technologies, integrations, and their resulting architecture. It is also imperative that buy-in from all stakeholders be obtained in order to ensure common understanding among all participants with regard to both what will be evaluated and in what manner.

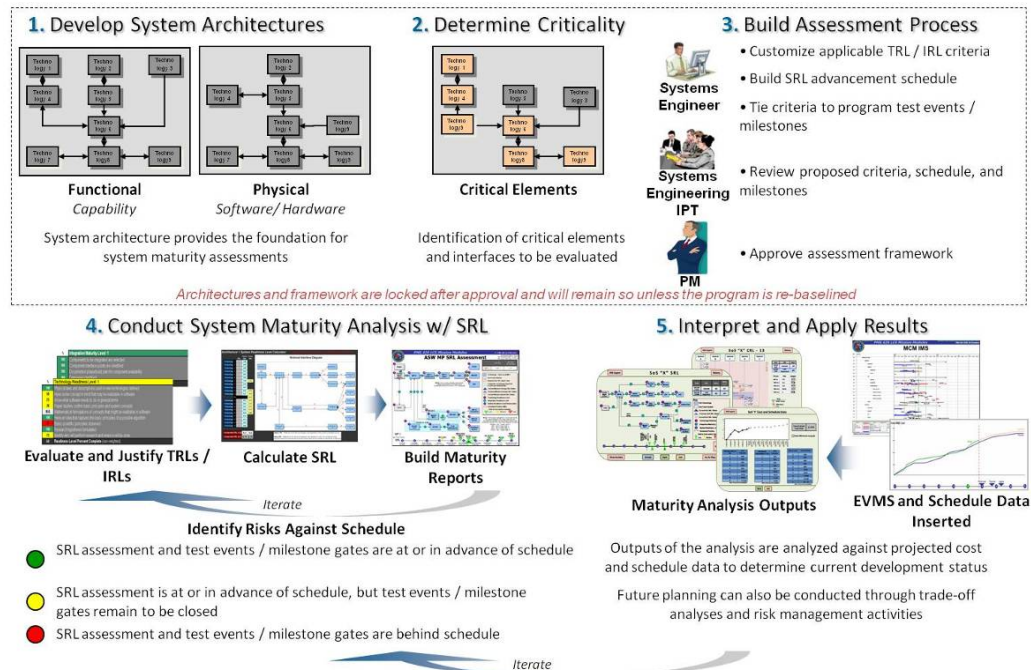


Figure 4. SMA Development Process Flow

Requirements Management & the Drive Towards Commonality in a SoS World

One of the more challenging of the SoS systems engineering tasks is the management of requirements. This is particularly acute in acknowledged SoS, where the PM does not have complete control of many of the constituent systems that are integrated into the SoS. For the program, this issue was made more acute by the initialization of the program on a “come as you are basis” based on the desire to explore the concept of modular MPs supporting a Seaframe using defined interfaces before fully investing in a full up acquisition program. This resulted in technologies being selected to satisfy the original performance requirements on a package-by-package basis. As the entire LCS program (Seaframes and MMs) developed, this concept of an experimental development morphed into a traditional acquisition program. The legacy of the initiation was a set of MPs with minimal commonality between the packages, as indicated on the left side of Figure 5. This approach, if continued, would result in increased LCCs for the support of the three MPs. While the way this issue was initiated for the LCS is slightly unique, the challenge of seeking to reduce LCC for a SoS in which the PM may not have direct control over the individual MSs and their designs is not.

The program is addressing this problem through the implementation of a structured and controlled process designed to introduce increasing levels of commonality across the various mission packages. The objective is to achieve a more flexible and controlled requirements and specification environment throughout the lifecycle development of MPs. To do this, the PMS 420 Systems Engineering Integrated Product Team (SE-IPT) established the objective of moving towards a target documentation structure, illustrated in

the right-hand side of Figure 5. This structure is designed to provide a migration path over time towards common capabilities. Initially, each package defined its separate capabilities with a package-level performance specification. Due to the initial compressed developmental schedule and the drive to use existing Program of Record solutions, duplicative or near duplicative common capabilities were often observed across the other MPs, as illustrated on the left-hand side of the figure.

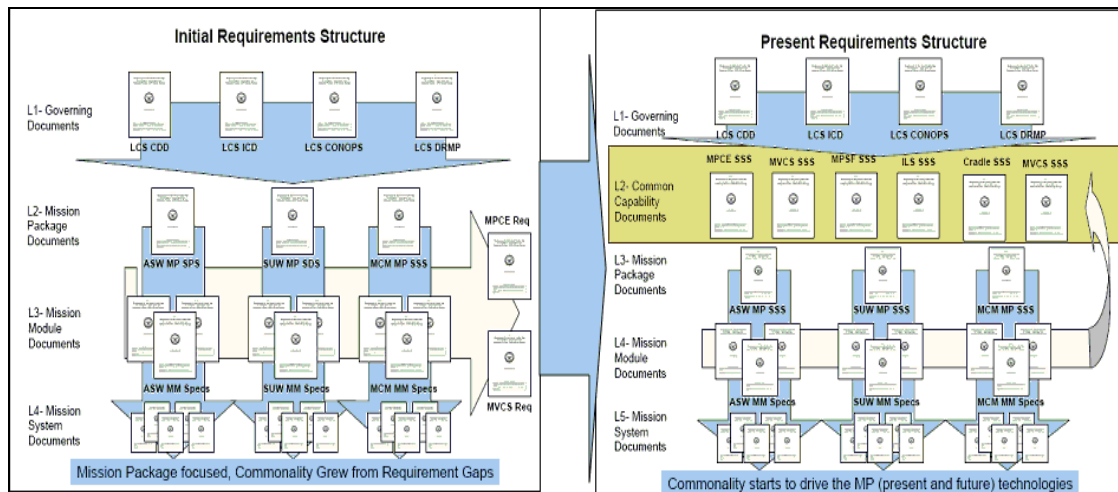


Figure 5. Modularizing the Structure of the Documentation Database

PMS 420, through the SE-IPT, has established a process to define commonality and modularize capabilities to support consistent re-use across MPs. In the first iteration of this process, several common MP components were isolated and established as a common set of requirements at the next level of specification below the governing requirements documentation. This adds a common layer that spans capabilities across individual MS boundaries, as illustrated on the right side of Figure 5, modularizing the designs and fostering re-use of capability. This objective was implemented early in package development, even while the three MPs were being developed under an accelerated schedule. While the limitations in technology and time prevented the implementation of desired common capabilities such as a common Unmanned Surface Vehicle (USV) between the MCM and ASW MPs, some commonality was achieved and implemented, including the design and use of a common USV cradle, common aviation support container designs, and other components.

In other areas, while the need for commonality across all three MPs was defined early, the ability to reach a common capability took time. These common capabilities included a modularized Mission Package Computing Environment (MPCE), and a common off-board unmanned vehicle communications capability, the Multi-Vehicle Control System (MVCS). Each of these common capabilities, i.e., MPCE has its own System/Subsystem Specification (SSS) and lower-level documentation. While PMS 420 designated these as common cross package capabilities to fulfill the identified needs of the three MPs, they have matured into requirements constraints on the existing packages. PMS 420 uses the development of a common set of interface requirements for defining the trade space for evaluating the value of incorporating new technologies within existing MPs.

This methodology, used for developing and transitioning MPCE and MVCS is now, through the use of an established requirements management and allocation process, being

extended to extracted common requirements and supporting capabilities in such areas as Integrated Logistics Support (ILS), Safety, Shock and Vibration, and others. As this method continues to mature, it will provide the basis for defining the totality of the interface requirements that a new mission package would need to met to be effectively and efficiently integrated into the LCS MP SoSs. The next phase beyond that will be to extend the push for commonality down to the MM and MS levels to include the identification, development, and specification of common MSs across the MPs, including shared capabilities such as remote unmanned vehicles. It is anticipated that this will allow for increased modularity and re-use of other Commercial Off The Shelf/Government Off The Shelf (COTS/GOTS) technology.

Steady progress has been made in reaching this planned requirements structure of the DOORS database. The requirements allocation analyses for the Flight 0 CDD, Flight 0+ CDD and the LCS Interface Control Document (ICD) have been completed and the Level 2 MPCE and MVCS SSS documents and the Level 3 MP SSS documents are in the final stages of PMS 420 Configuration Control Board (CCB) approval. Formal configuration management (CM) process are used in updating and tracing the SSS documents to the CDD and ICD allocated requirements, which provides impact and traceability analysis capability to the MP level.

Cost Prediction and Monitoring of SoS

PMS 420 has been developing a SoS Online PM Tool that significantly improves the ability to manage the program. Its advantages are numerous, including:

- a. Earned Value Management at all SoS levels, including the ability to examine the data across various cross-sections of the program, and forecast cost and schedule overruns at an early stage;
- b. Support of multi-year planning of research, development, procurement, and sustainment efforts;
- c. Risk management that associates risk with cost and impact;
- d. Integrates a formal program change management process across all levels of the program;
- e. Allows senior program management the ability to conduct what-if financial impact analysis without changing the baseline [what's the impact if system (A) is replaced with system (B)];
- f. Links the Integrated Master Schedule (IMS) to task planning and task execution both at the system level and the SoS level;
- g. Links the Acquisition Program Baseline (APB) to execution data. This gives the program manager critical insight into how lower-level system performance will impact the overall success of the SoS; and
- h. Integration of the System Maturity Model into the Earned Value Management System (EVMS).

PMS 420 uses this tool in its monthly drumbeat process, shown in Figure 6. The web-accessible, CAC-enabled Online PM Tool integrates program planning and execution data (cost, schedule, and performance), ensuring proper visibility into who is doing what, how well it is being done and what has to be done next. Detailed task planning and execution data, coupled with clever organizational and reporting capabilities, have allowed the PM to connect program priorities and goals with past EV performance and future EV



forecasting. The Online PM Tool provides the framework and business rules to establish and maintain process discipline. PMS 420, executing activities, and cross-functional stakeholders all have a “seat at the table.” The real-time data availability promotes two-way communication and concurrence, enabling a successful egalitarian approach that would not be possible without the open web-based tool environment. This provides the PM with insight into potential challenges and has helped avoid and/or minimize program errors.

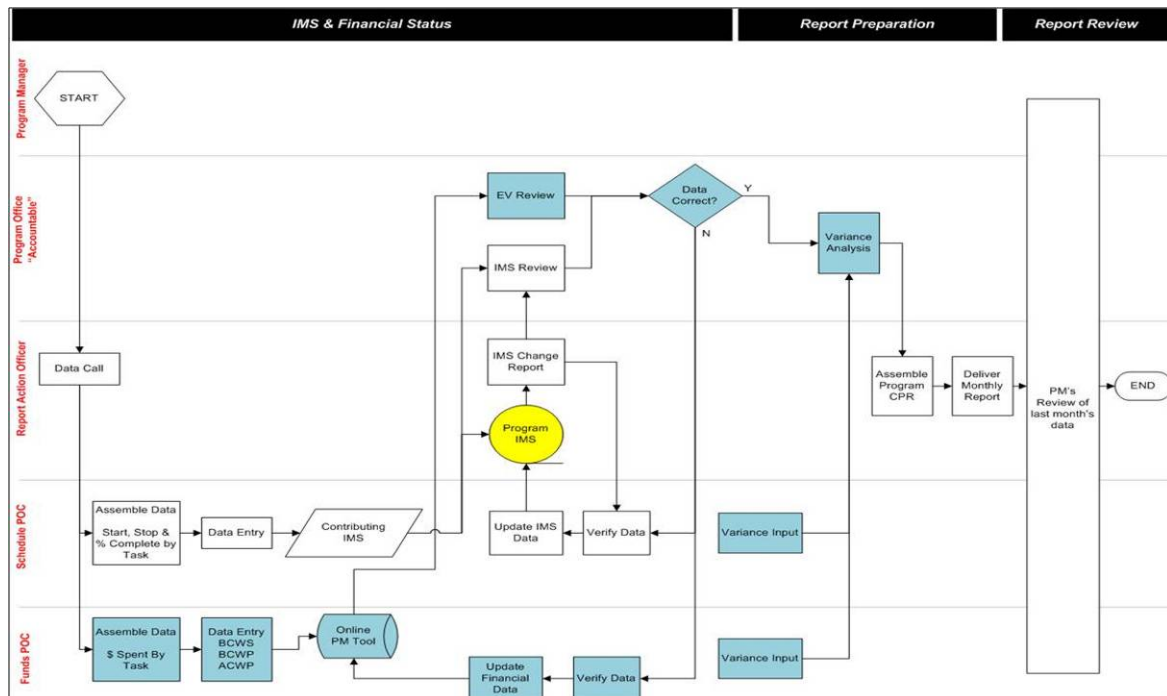


Figure 6. Monthly Drumbeat Process, Work Performance Insight and Reporting

The Online PM Tool has been developed using a process that incorporates key stakeholder requirements. Processes are mapped to distinct decision points and controls implemented to ensure results are achievable. The development team used an incremental build methodology to develop the Work Breakdown Structure (WBS), record processes, review tool development and obtain user feedback. The PMS 420 WBS is used as the unifying element for program information and all information in the Online PM Tool is associated with a specific WBS element. Individual task statements are assigned a specific WBS element and funding is then allocated to executing activities associated with the specific task statements. Execution data (earned value data and variance analyses) is reported against specific WBS elements, which are then displayed in EV reports. The detailed use of EV and detailed work plans identifies areas that require closer scrutiny, and have provided the PM with increased insight into work performance, as shown in Figure 7. Areas with excess or un-executable funds are also more easily identifiable, allowing those funds to be reallocated to other program priorities, as required.

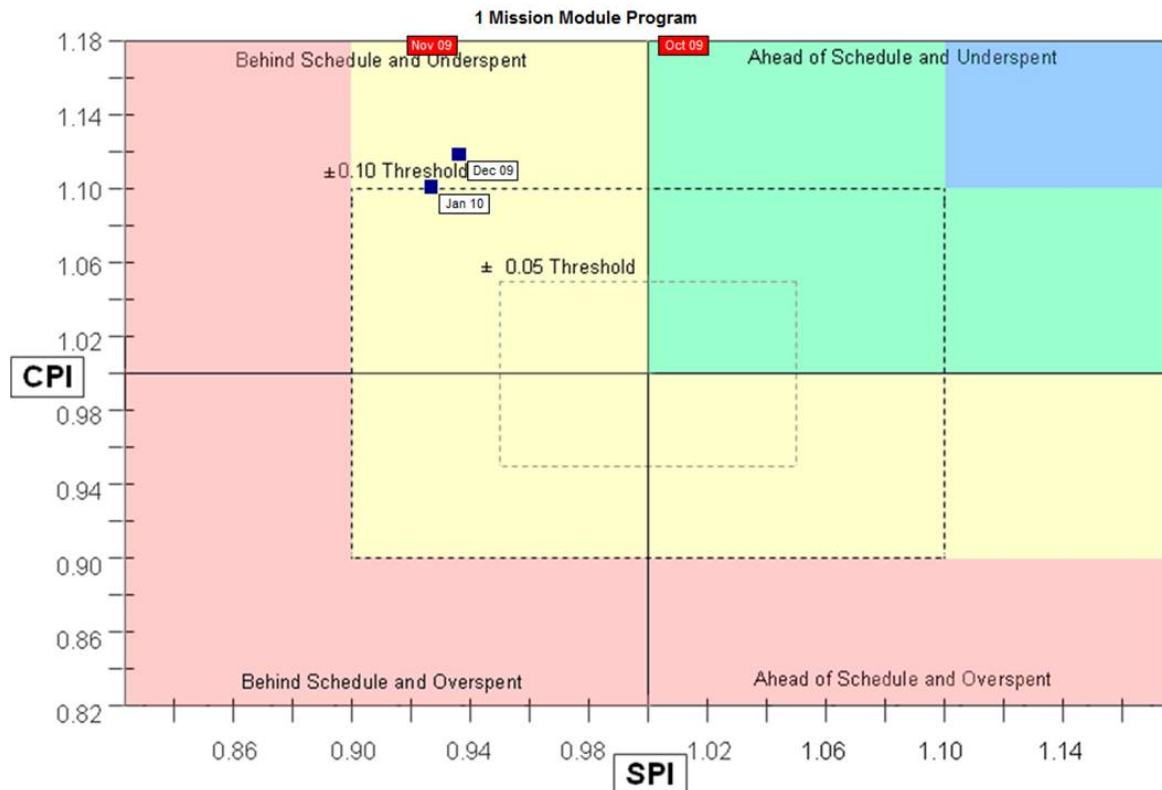


Figure 7. Work Performance Insight

PMS 420 uses the Online PM Tool to display, discuss and digest data during monthly drumbeat meetings. This has enabled the monthly drumbeat to evolve from a tool-centric execution status meeting to a more generalized program management meeting where Cost, Schedule, Risk and Performance are reviewed. This meeting has come to replace other routine meetings, concentrating PM efforts into a single day and freeing program office staff to do other work. Snapshot documents, i.e., program Cost Performance Report and Integrated Master Schedule, are created using the tool to summarize and synthesize data into information, record this information for historical purposes, and provide senior management with real-time status of program health.

PEO-level reporting is accomplished through a PEO Dashboard, which provides senior executive-level insight into all programs in the PEO. Senior management has the ability to view planning and execution performance data across program offices, executing and performing activities, contractors, enterprise mission capabilities, lifecycle stages, and across SoS and Families of Systems. PEO Quarterly Execution Reviews (QER), previously an arduous process of data collection and validation, are accomplished through the use of the PEO Dashboard and Online PM Tool. The PM can generate a PEO-level summary of the same data used to manage the program, and the PEO can request deep-dives into areas of interest or concern.

Understanding and Influencing SoS Reliability

Availability is a more complex problem for the SoS than in traditional systems. This arises primarily out of two attributes of many SoSs; the use of modular systems design, and the ability of the SoS to accomplish multiple mission capabilities using various components of the SoS. The program has been addressing this availability issue through several

analysis approaches and the development of tailored methods linked to the planned operational concepts of the MP SoSs.

The use of traditional availability and reliability tools, such as reliability block diagrams (RBDs), Failure Mode Evaluation and Criticality Analysis (FMECA), and others are well understood and there is much experience in industry with their application. However, in a SoS, these methods are faced with two added demands. First, the introduction of modularity, open system design, and remote operations produce increasing number of components and, therefore, more opportunity for component to component failures. Some of the impacts on operational availability in such extended systems are indicated in the list below:

- The mission component string is inherently less reliable because we increase the number of serial components in the mission/operational function;
- Extended Unmanned systems require set up time and the potential for damage is increased because of the increased handling, in addition the deployment and recovery environment and handling systems design introduce opportunities for damage;
- Infrastructure Over-head can be over whelming in the particular adaptation of modular Plug and Play (P&P) design approach (weight, extra services, handling operations, software and hardware overhead); and
- Deployment of remote systems have security challenges (physical and data related).

The second added demand is that of multiple mission capabilities and flexibility in configurations to achieve them. Simply stated, we typically have more capabilities provided by the SoS, and can execute the capability with several combinations of components. Some of the challenges introduced by this flexibility are described in the list below:

- Operational use is constrained during a specified time period;
- Operational use may use a small percentage of the mission suite, depending on the mission, e.g., for MCM, mapping, identification, clearing; and
- Operational environment may call for a different subset of equipment to be used in a deployment.

In the case of the program, we have the following specific issues that we must address:

- The majority of the elements of a MP are different and widely distributed interfaces will affect the availability;
- The organic off board vehicles can be utilized differently each mission;
- The utilization of mission equipment per mission will vary—alternative mission equipment may be substituted;
- The deployment and utility times will vary based on operational mission goals;
- Ship Availability is a component of the MP architecture; and
- Software reliability in MPCE and MVCS needs to be considered.

An additional issue for the program (and becoming prevalent in other SoS) is that the requirement for MP availability is defined in terms of a Materiel Availability KPP. Materiel Availability (A_m) for the LCS MPs is established with a threshold of 0.64 and an objective of



0.712. There is no specified separate requirement for the traditional A_o , since the LCS CDD indicates that “it is embedded in the Materiel Availability KPP.” A novel methodology has been applied to analyze and decompose A_m , resulting in an approach that separates A_m into two components. The first, called Active Availability (A_a), is a factor that is a function of fleet-level support design, including such components as depot-level repair requirements, fielding, deployment, and support strategies. The second is the traditional A_o , defined at the fleet level and computed as the average of squadron or unit level A_o .

In this fashion, a clear definition of A_m for the program is developed and represented by $A_m = A_a * A_o$. Monte Carlo simulations and support parameter investigations are being used to determine the impact on MP A_m . Initial results indicate that target parameters for A_a and A_o with their impact on A_m could be established as given in the example in Table 1.

Table 1. Example LCS MP A_m Composition

CDD Requirement	Target A_a	Target A_o	Target A_m
Threshold .64	0.887	0.738	0.655
Objective .712	0.887	0.809	0.718

Using this A_m decomposition and simulation method, it is possible to establish requirements for both A_a and A_o that can be allocated to the MPs and further decomposed to their individual MM and MS components. The primary use of the A_m decomposition method is to establish support concepts and strategies that will establish a defined level of A_a . After A_a is clearly established, A_o can be determined using the equation $A_o = A_m / A_a$. This provides target A_o requirements that can be allocated to the MPs and their components. To address the above issues, Operational component strings for the various mission functions are defined. An example of this is shown in Figure 8 below.

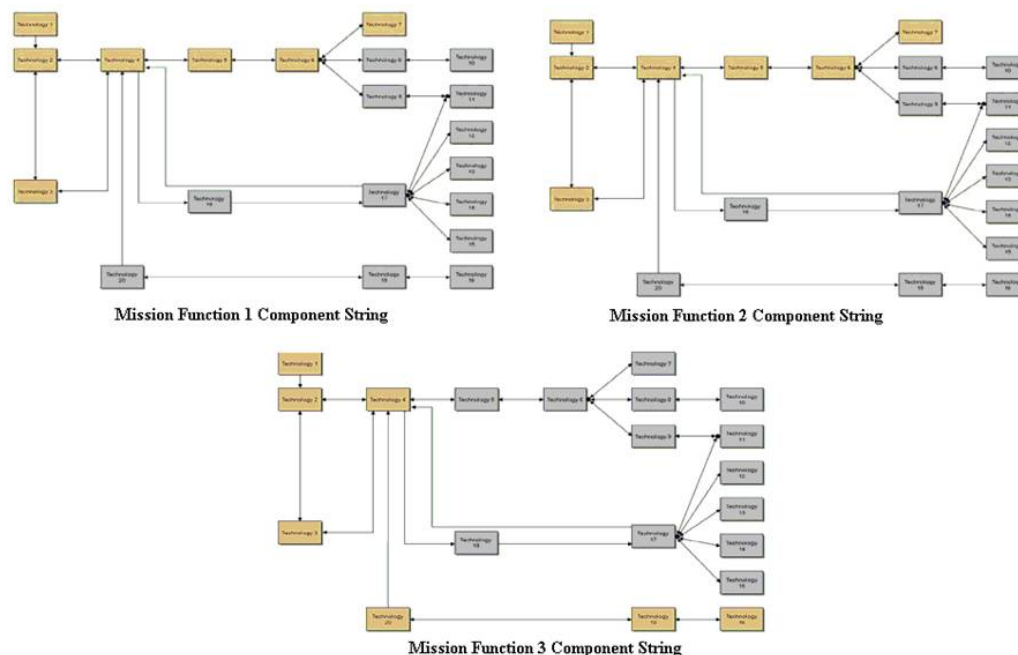


Figure 8. Mission Function Component Strings

The following steps are then performed on the individual mission strings:

- Operational strings were analyzed to identify the components required to execute independent mission functions of the system;
- An assessment of the string to achieve a Mission A_o contribution is performed;
- Common components (nodes) that form a critical function in more than one mission function are identified, and operational time is calculated for each mission it touches over the deployment cycle; and
- Allocation of the Mission A_o decomposes to an A_o requirement at the component, Line Replaceable Unit (LRU), level.

Currently, the program is evaluating the use of the mission function strings, decomposition of the A_o for the mission to individual components, and a Monte Carlo simulation for mission A_o analysis. The goal is to determine individual component A_o requirement targets and determine the impact of these on MP A_o and, therefore, on the A_m threshold and objective requirements.

Expanding the Tool Set—Technology Trades & SoS Performance Predictions

The existing tool set has been invaluable to enabling the PM to gain increased insights into the developmental status and risks associated within a complex SoS. PMS 420 has now begun working to expand upon the foundation of those tools and the system readiness monitoring capabilities provided by implementation of the SMM methodology and is seeking to expand the technical and programmatic tools and methodologies available for the PM. Specifically, there exists the need for tools to assist the PM in understanding the impacts of technology insertion options and to gain insight into predicted SoS performance, enabling the more effective conduct of tradeoff analysis. This analysis would assess SoS performance and capability objectives and provide recommendations enabling the PM to make choices that optimize the SoS on the basis of cost, technical risk, or anticipated performance.

Technology Insertion & SoS Analysis

One of the primary benefits of the modular approach to SoS development is that new technologies capabilities can be rapidly incorporated to improve reliability, performance, or reduce LCCs. However, comparable technologies when integrated may result in significantly different integration risks, performance impacts across the SoS, and reliability or cost impacts. How to decide on what technologies to change and which technologies to select is an area of critical interest to PMS 420 as the MPs mature from development and enter their operational lifecycles.

As the SoS technologies reach their designed or actual limits (of cost, performance, etc.), they should be reviewed for replacement by newer or more robust technologies that would provide improvement to the SoS optimization, performance, and capabilities. The existing methodologies and incorporation of tools developed to date by PMS 420 have provided the program with an unprecedented view into the details of the status of the SoSs. The difficulty lies with combining the various details to provide the PM with a current composite view into the SoS to support analysis and impacts related to changing technologies within the existing SoS. A composite view would enable the PM to determine if a proposed Technology would generate performance beneficial or detrimental to the SoS,



would have a budget that is proportional to established values—that is, reaching the end of its lifecycle— would bring too much risk to the SoS, would be within the physical constraints (size, mass, etc.) of the SoS, and more.

The question is how to review several technologies that could be added to the SoS and determine the best candidate based upon the desires of the PM. To make this decision consistently and without bias, a modified version of the Analytical Hierarchical Process (AHP) will be utilized. With AHP, PMS 420 can assign several categories and sub-categories that the various technologies will be rated against, such as Cost, Physical Constraints, SRL, Reliability, etc. The scores from the categorical comparison of the technology ratings will then be weighted by the needs and desires of the PM (e.g., budgetary constraints). The result will be a list of the current and potential technologies ranked in order of recommended choice. A basic example of this hierarchical calculation is shown in Figure 9.

Technology Analysis and Insertion Tool Development

The difficulty with utilizing the technology analysis and insertion calculation is its complexity and the need to allow for easy modification to the weighting parameters and ratings of technologies. In the complex world of program acquisition, there are times that this calculation will need to be finished in a relatively short period of time, demanding an implementation that is prompt, dependable, unbiased, and accurate. Much of the data that is needed for the calculation is spread out in many tools and locations and would need to be accessed for the calculation to produce meaningful results.

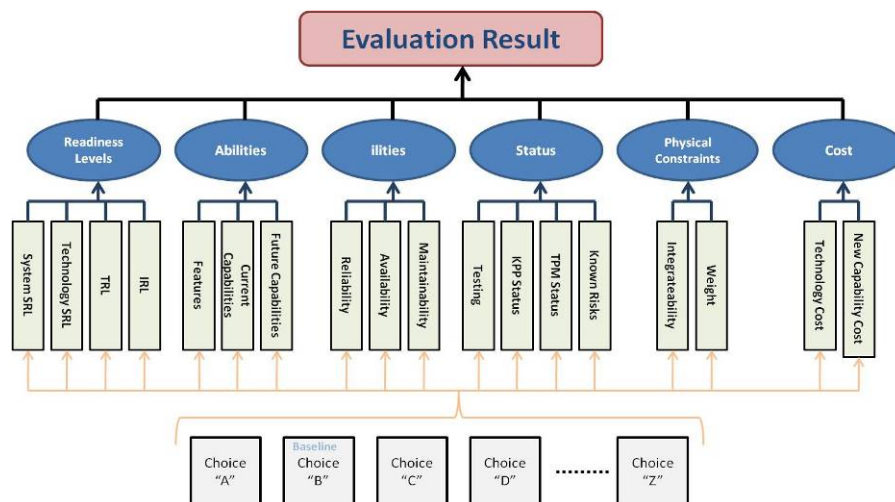


Figure 9. Technology Analysis Hierarchy

The solution to these issues is to maximize the usage of the current tools while minimizing the footprint of a new one. We have been working with the SIT to develop and test a SRL calculation plug-in for an architecture tool, as a distributable example. Building off of this starting point, we have designed this trade-off tool to act as a substantial plug-in to the program's already established architecture modeling tool. By leveraging the SoS architecture that is already developed and being maintained and by adding small changes to the architecture tool's data model, the new trade-off tool will function with the established architecture tool, greatly reducing the learning curve.

The small modifications to the architecture tool data model will not impact any of the existing programmatic data that is currently captured, nor will it impact the normal

architecture review cycles. (As the data model can easily be expanded to incorporate the new fields, it can be designed to fit into DODAF 1.x and 2.0 versions.) The modified fields will capture the ranking data that relates to the Technology Analysis and Insertion calculation in the configuration controlled environment of the Architecture Model. The information for technologies that are not currently part of the SoS would be populated into the architecture tool. A small modification to SIT's existing SRL calculation plug-in, will enable it to work in this design (different architecture tool's commands). The weighting values for the criteria will be entered by the PM and securely stored. The outcome of the calculation will be a ranked list of choices for the position within the SoS. A basic representation of the tool is shown in Figure 10.

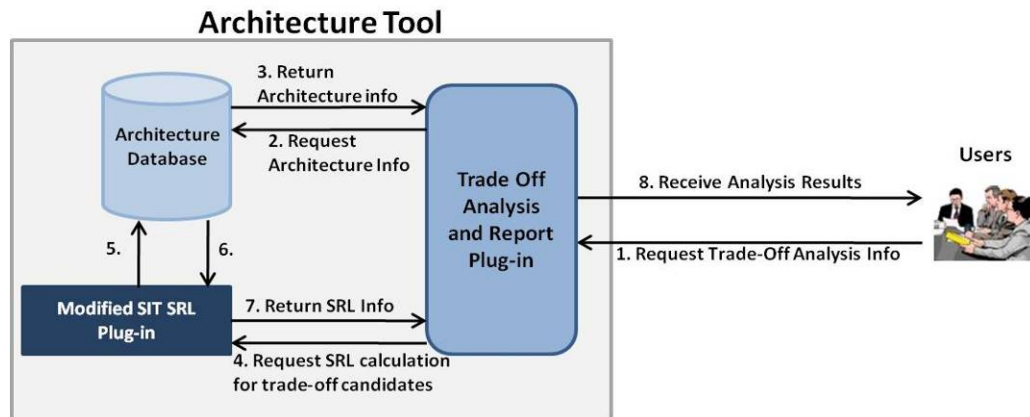


Figure 10. Technology Analysis and Insertion Tool

Prediction of Performance Using a Performance Level Monitoring Methodology

Through system development, PMs are expected to quantifiably justify that their program will result in the delivery of a system with the required performance. The traditional PM has several technical and program management tools at their disposal, including TPMs, Modeling and Simulation, etc., that provide insight and predictive capability in system performance. When the program matures to a point at which actual test data can be gathered, it is compared against expected system performance. Due to the complex nature of SoS interdependencies, PMs are especially challenged when asked to quantifiably justify the investment in time, personnel, financial, and material resources in the program during SoS development.

Traditional program and technical management tools must be extended to provide the necessary insight to the acknowledged SoS environment. Given that the performance of a SoS is directly dependent upon the performance levels of the individual systems composing the SoS, as these capabilities are being independently developed by PARMS (over which he has limited directional authority and who may be developing the capabilities to fulfill a different set of performance metrics or may be unwilling to share detailed technical status with external organizations). Even where the individual MS performance may directly translate to a SoS KPP area, the nature of MP SoS and its Concept of Operations does not mean that it provides the total answer. In various scenarios, the individual MP KPPs can be achieved through using various combinations of the systems within the SoS. For example, the LCS may decide to engage Fast Inshore Attack Craft (FIAC) with either the LCS's core gun, capabilities onboard the MH-60 helicopter, the 30mm Gun Mission Module (GMM) (of

the SUW MP), or eventually the missiles of the Surface to Surface Missile MM (of the SUW MP). All of which can in full or in combination satisfy individual KPPs for the SUW MP. This estimation of performance difficulty is further complicated when the SoS is being developed in an incremental manner. Again using the SUW MP as an example, the first SUW MP is currently installed onboard LCS-1 and includes two 30-mm Gun Mission Modules, an Aviation (MH-60R armed helicopter) MM, and a prototype MSM. The first SUW MP does not include the Surface-to-Surface Missile MM (based on the NLOS-LS), which will be added in Increment 2. Understanding the capability provided by MP increments and ultimately whether the baseline (full capability) MP will satisfy the full set of performance requirements is of interest to the PM. A final complication is that when conducting SoS and incremental acquisition of this nature, complete E2E test and evaluation (T&E) may not be feasible and computer intensive modeling and simulation may not be practical in a schedule driven environment or where the SoS PM may not have the full technical models of the individual systems. So within these limitations, how can the PM gain insight into, and predictive capability for, determining the ability of the SoS to achieve required performance?

To answer these questions, PMS 420, in conjunction with SSC-Pacific, Northrop Grumman, and the SIT, has been expanding the SMM to incorporate a Performance Level Monitoring (PLM) methodology. The PLM is being developed to understand if the performance will satisfy the KPPs and to understand the deltas in performance between the initial MPs and later MP increments, which will provide the full-up MP capability. Ultimately, this tool will also support the analysis of mission threads using different MP configurations, i.e., providing insight in performance capability of the MCM MP if one of its USVs is down for maintenance and/or as a tool for evaluating the impact of incorporating new capabilities or changing existing capabilities within a MP.

Performance Level Monitoring Explained

The PLM strives to apply a modified TPM type approach to a SoS construct. However, instead of focusing on a measurable technical value that can be monitored during development within a individual system, the PML links the SoS KPPs to individual component capabilities, their maturity, and their potential usage. The SMM, Concept of Operations (CONOPS), and usage rate variance analyses are all considered in the PLM calculation.

To implement this process, significant up-front evaluation will be required by the SoS program office. The first step of the methodology is to define the SoS MP in terms of its component MSs and to map those systems and their capabilities to their projected impact on satisfying the MP KPPs. The individual MS capabilities are then adjudicated by the SoS PM in terms of their maturity and inclusion within an individual MP, breaking them into three generic categories of Advanced Developmental Models (ADM), Engineering Development Modules (EDM), and Production Models (PROD) that are mapped to their expected maturation points over the analysis timeline. This adjudication of systems is represented through the use of a weighting function to represent the individual capability's maturity (real or anticipated) for each level of development. While this method works well for MS that are not integrated with others to deliver required capability, such as the GMM, this becomes more complex when two or more systems must come together to provide a level of capability such as a MM, for example the combination of the ASW USV MS and the USV Towed Array System (UTAS) to provide a passive search capability. Fortunately, the ongoing development of the SMM concept allows for a potential approach by using the value calculated for the MM SRL. The individual technologies can then be weighted in terms of their contribution to the accomplishment of the capability and be combined into a series of



capabilities or MM values. The integrated MM capabilities can be expressed as a single value where the level of capability that the module comprised of capabilities (x, y, and ..) that can contribute towards the satisfaction of the MP KPP requirement given the level of maturity of the capability in the MP.

The next stage of the PLM is to define the impact of various CONOPS on the ability of the SoS to satisfy the KPP requirements. As one of the strengths of a SoS is its inherent flexibility where component systems can be organized to solve the capability problem, this can often be translated to where the individual capabilities may be used in varying ways to accomplish the same mission. These varying CONOPS result in increased complexity for the analyst in trying to predict what level of performance is being achieved. While modeling and simulation tools can be used to conduct this type of analysis, the variations would make this an expensive and time-consuming process for the program and would not provide the PM with the rapid insight into options that maybe required. To address this issue, the PLM seeks to develop a set of scenarios for each of the MP SoS that represent the range of potential operational usage concepts for the individual capabilities (or modules) within the SoS as applicable to each KPP and incremental MP. This enables the derivation of a set of equations relating the KPPs to their component technologies and to a specific CONOP. The set of CONOPS, each reflecting an anticipated level of performance and technical maturity/integration of a specific capability (x) at a specific point (for this example as represented by a specific mission package) in time (n) are then matrixed together to enable a calculation of the overall predicated performance of the SoS across a range of scenarios. When conducting the analysis, the exact usage rate for each of the MS/MM may be unknown. In this case, running the analysis for each CONOPS using a minimum, maximum, and average anticipated usage rate for each capability/module can be used to develop a set of error bars in performance predictions. As a tool for the management of risk and for predicting when performance will be achieved or to understand the potential impact of changes, a graphic similar to the traditional TPM graphic can then be constructed by calculating composite KPP values for each MP increment and plotting the composite level of performance against time.

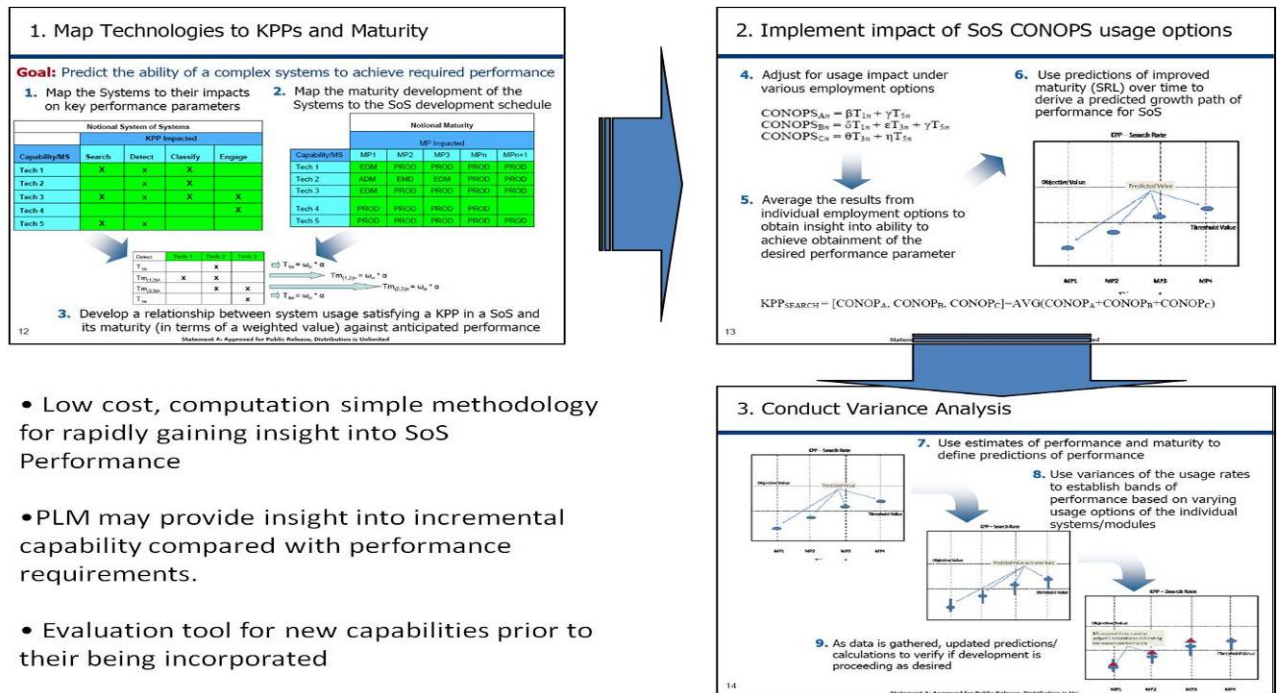


Figure 11. Performance Level Assessment Methodology

The intent behind the PLM is to provide the PM necessary insight into incremental capability compared with CDD performance requirements. As with all predictive models, the analysis will need to be compared against measured test data as it becomes available to verify predictions and identify if the program is on course to meet CDD performance requirements. Figure 11 provides a graphical flow of the methodology described above and a more detailed description of this methodology can be found in *Notional Assessment Methodology for KPP Accomplishment in a SoS: Proposed Methodology for Measuring Performance Progress within a System of Systems (SoS)* (Volkert, 2009). This methodology has been further expanded upon by SIT at this symposium (Tan, Sauser, Ramirez-Marquez, Magnaye, Nowicki & Deshmukh, 2010). A projected further expansion of this methodology will be to allow for the evaluation of new capabilities prior to their being incorporated into a planned upgrade or replacement of an existing capability.

Conclusion

The increasing use of the System of Systems (SoS) model for the fielding of new and improved warfighting capabilities poses new management challenges for the DoD. To support the Program Manager, the US Navy (PMS 420, SSC Pacific), Northrop Grumman, and the Stevens Institute of Technology have been collaborating on the development and verification of a set of comprehensive financial and portfolio management methodologies for acknowledged SoSs. The tools and capabilities that are being developed, discussed, and expanded by PMS 420 reflect the real-world challenges facing a SoS PM and reflect valuable lessons learned to date within the LCS Mission Modules program. Starting with the field of technology maturation, the team has developed the standard TRL methodology into a concept that develops a System Readiness Level (SRL) measurement as a measure of SoS integration and maturity. This methodology has been demonstrated and has been used as the developmental springboard into an approach for determining and predicting the probability of achieving system performance and for understanding the impact of technical

option trades. Financial tools have been developed and implemented that allow the PM to gain insight beyond that afforded by traditional EV reporting and that can provide management assistance in resource allocation in dynamic programs. The maturation of the requirements management process for a SoS was discussed and the capabilities of using it as a tool for reducing Life Cycle Cost presented. The process described dictates the need for a methodology for SoSs that allows for fully integrated analysis and trade-offs of the technical, cost, and schedule design spaces. While SoS show great promise for providing flexible and cost-effective provisioning of capabilities to the DoD, the evolution of management tools will need to continue to advance in order to allow for the more efficient application of scarce resources from the conception of program initiation. Otherwise, SoS Program managers may be forced to continue to face many of the challenges PMS 420 has been through and will need to expend resources in solving those management challenges vice applying the resources to product development.

References

- Forbes, E., Volkert, R., Gentile, P., & Michaud, K. (2009). Implementation of a methodology supporting a comprehensive systems-of-systems maturity analysis for use by the littoral combat ship mission module program. In *Proceedings of the Sixth Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School.
- GAO. (2006, September 14). *Best practices: Stronger practices needed to improve DOD technology transition processes* (GAO-06-883). Washington, DC: Author.
- PMS 420, Littoral and Mine Warfare, US Navy. (2009). *System maturity assessment guide*. Scientific and Technical Reports, Littoral Combat Ship Mission Package Program. Washington, DC Navy Yard.
- ODUSD(A&T)SSE. (2008). *Systems engineering guide for systems of systems* (Vers. 1.0). Washington, DC: Author.
- Sauser, B.J., Ramirez-Marquez, J.E., Magnaye, R.B., & Tan, W. (2008). System maturity indices for decision support in the defense acquisition process. In *Proceedings of the Fifth Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School.
- Tan, W., Sauser, B., Ramirez-Marquez, J., Magnaye, R., Nowicki, D., & Deshmukh, A. (2010). System capability satisficing in defense acquisition via element importance measures. In *Proceedings of the Seventh Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School.
- Volkert. (2009, September). *Notional assessment methodology for KPP accomplishment in a SoS: Proposed methodology for measuring performance progress within a system of systems* (SoS) (PMS 420 White Paper).



2003 - 2010 Sponsored Research Topics

Acquisition Management

- Acquiring Combat Capability via Public-Private Partnerships (PPPs)
- BCA: Contractor vs. Organic Growth
- Defense Industry Consolidation
- EU-US Defense Industrial Relationships
- Knowledge Value Added (KVA) + Real Options (RO) Applied to Shipyard Planning Processes
- Managing the Services Supply Chain
- MOSA Contracting Implications
- Portfolio Optimization via KVA + RO
- Private Military Sector
- Software Requirements for OA
- Spiral Development
- Strategy for Defense Acquisition Research
- The Software, Hardware Asset Reuse Enterprise (SHARE) repository

Contract Management

- Commodity Sourcing Strategies
- Contracting Government Procurement Functions
- Contractors in 21st-century Combat Zone
- Joint Contingency Contracting
- Model for Optimizing Contingency Contracting, Planning and Execution
- Navy Contract Writing Guide
- Past Performance in Source Selection
- Strategic Contingency Contracting
- Transforming DoD Contract Closeout
- USAF Energy Savings Performance Contracts
- USAF IT Commodity Council
- USMC Contingency Contracting

Financial Management

- Acquisitions via Leasing: MPS case
- Budget Scoring
- Budgeting for Capabilities-based Planning



- Capital Budgeting for the DoD
- Energy Saving Contracts/DoD Mobile Assets
- Financing DoD Budget via PPPs
- Lessons from Private Sector Capital Budgeting for DoD Acquisition Budgeting Reform
- PPPs and Government Financing
- ROI of Information Warfare Systems
- Special Termination Liability in MDAPs
- Strategic Sourcing
- Transaction Cost Economics (TCE) to Improve Cost Estimates

Human Resources

- Indefinite Reenlistment
- Individual Augmentation
- Learning Management Systems
- Moral Conduct Waivers and First-term Attrition
- Retention
- The Navy's Selective Reenlistment Bonus (SRB) Management System
- Tuition Assistance

Logistics Management

- Analysis of LAV Depot Maintenance
- Army LOG MOD
- ASDS Product Support Analysis
- Cold-chain Logistics
- Contractors Supporting Military Operations
- Diffusion/Variability on Vendor Performance Evaluation
- Evolutionary Acquisition
- Lean Six Sigma to Reduce Costs and Improve Readiness
- Naval Aviation Maintenance and Process Improvement (2)
- Optimizing CIWS Lifecycle Support (LCS)
- Outsourcing the Pearl Harbor MK-48 Intermediate Maintenance Activity
- Pallet Management System
- PBL (4)
- Privatization-NOSL/NAWCI
- RFID (6)



- Risk Analysis for Performance-based Logistics
- R-TOC AEGIS Microwave Power Tubes
- Sense-and-Respond Logistics Network
- Strategic Sourcing

Program Management

- Building Collaborative Capacity
- Business Process Reengineering (BPR) for LCS Mission Module Acquisition
- Collaborative IT Tools Leveraging Competence
- Contractor vs. Organic Support
- Knowledge, Responsibilities and Decision Rights in MDAPs
- KVA Applied to AEGIS and SSDS
- Managing the Service Supply Chain
- Measuring Uncertainty in Earned Value
- Organizational Modeling and Simulation
- Public-Private Partnership
- Terminating Your Own Program
- Utilizing Collaborative and Three-dimensional Imaging Technology

A complete listing and electronic copies of published research are available on our website:
www.acquisitionresearch.org





ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL
555 DYER ROAD, INGERSOLL HALL
MONTEREY, CALIFORNIA 93943

www.acquisitionresearch.org